```
!!SEQUENCE_LIST 1.0
! FINDPATTERNS on geneseqp:* allowing 0 mismatches
!       1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C,T,S)(R,K,H
    GENESEQP:P90009  ck: 7203  len: 102   finds: 1   ! Mouse gamma-2b chain New immun
    GENESEQP:P91884  ck: 718   len: 48    finds: 1   ! Antigenic Epstein-Barr virus p
    GENESEQP:P94157  ck: 2829  len: 246   finds: 1   ! Acetyl-CoA reductase. Construc
    GENESEQP:P90460  ck: 560   len: 288   finds: 1   ! Alpha-factor profibrolase from
    GENESEQP:P81999  ck: 4995  len: 25    finds: 1   ! Binding peptide 2 capable of h
    GENESEQP:P70477  ck: 9828  len: 194   finds: 1   ! Sequence of human respiratory
    GENESEQP:P70785  ck: 9828  len: 194   finds: 1   ! Sequence encoding human respir
    GENESEQP:R24296  ck: 3600  len: 384   finds: 1   ! Regulatory protein VanS involv
    GENESEQP:R10995  ck: 1620  len: 398   finds: 1   ! Xenopus Bone Morphogenetic Fac
    GENESEQP:R24190  ck: 5872  len: 186   finds: 1   ! Bovine RSV strain A 51908 M2 p
    GENESEQP:P81092  ck: 3060  len: 95    finds: 1   ! Sequence of rhinovirus HRV2 vi
    GENESEQP:P81093  ck: 5281  len: 323   finds: 1   ! Sequence of rhinovirus HRV2 vi
    GENESEQP:R10974  ck: 3836  len: 246   finds: 1   ! Acetoacetyl CoA reductase enzy
    GENESEQP:R12468  ck: 7673  len: 248   finds: 1   ! Luffa cylindrica bioactive pro
    GENESEQP:R13252  ck: 528   len: 247   finds: 1   ! Murine Cytotoxic Cell Protease
    GENESEQP:P60645  ck: 2381  len: 292   finds: 1   ! Mouse kidney cell Band 3-like
    GENESEQP:R14247  ck: 3078  len: 149   finds: 1   ! HCV-T (6110-6557) encoded epit
    GENESEQP:R20004  ck: 4995  len: 25    finds: 1   ! VP6 inner capsid epitope, 84 T
    GENESEQP:R21409  ck: 617   len: 482   finds: 1   ! NADH dehydrogenase subunit 2.
    GENESEQP:R23004  ck: 7059  len: 297   finds: 1   ! Protein transcribed from the m
    GENESEQP:R22695  ck: 4919  len: 286   finds: 1   ! blaC31 mutation (5). Novel DNA
    GENESEQP:R25601  ck: 5700  len: 85    finds: 1   ! NANBH virus antigen. Antigenic
    GENESEQP:R25869  ck: 3783  len: 239   finds: 1   ! HCV polypeptide 16. Hepatitis
    GENESEQP:R25891  ck: 4644  len: 253   finds: 1   ! HK16. Hepatitis C virus antige
    GENESEQP:R25303  ck: 9668  len: 194   finds: 1   ! HRSV 22K protein. Vaccines for
    GENESEQP:R29874  ck: 9846  len: 171   finds: 1   ! HCV NS4-NS5 peptide N29-1, N29
    GENESEQP:R29879  ck: 8945  len: 285   finds: 1   ! HCV NS4-NS5 peptide 2918. New
    GENESEQP:R29895  ck: 8862  len: 171   finds: 1   ! HCV NS4-NS5 peptide N29-1. New
    GENESEQP:R29896  ck: 8862  len: 171   finds: 1   ! HCV NS4-NS5 peptide N29-2. New
    GENESEQP:R29897  ck: 2886  len: 167   finds: 1   ! HCV NS4-NS5 peptide N29-3. New
    GENESEQP:R29909  ck: 429   len: 277   finds: 1   ! Prod. of the luffin-f gene. Ne
    GENESEQP:R29910  ck: 9157  len: 278   finds: 1   ! Prod. of the luffin-g gene. Ne
    GENESEQP:R31608  ck: 7012  len: 205   finds: 1   ! Homologous to chicken nov gene
    GENESEQP:R30618  ck: 7998  len: 106   finds: 1   ! Polypeptide coded by Korean HC
    GENESEQP:R32192  ck: 2829  len: 246   finds: 1   ! Sequence encoded by the ace
    GENESEQP:R37294  ck: 7673  len: 248   finds: 1   ! Plant type I RIP Luffin. An
    GENESEQP:R39259  ck: 8652  len: 391   finds: 1   ! Human somatostatin receptor
    GENESEQP:R39260  ck: 8110  len: 391   finds: 1   ! Murine somatostatin recepto
    GENESEQP:R42317  ck: 3489  len: 345   finds: 1   ! EBV VCA-p40 from ORF BdRF1.
    GENESEQP:R52667  ck: 5723  len: 303   finds: 1   ! Equine herpesvirus US2 gene
    GENESEQP:R49296  ck: 4901  len: 25    finds: 1   ! Invarient chain (Ii) positi
    GENESEQP:R49297  ck: 3126  len: 24    finds: 1   ! Invarient chain (Ii) positi
    GENESEQP:R49299  ck: 1278  len: 23    finds: 1   ! Invarient chain (Ii) positi
    GENESEQP:R49438  ck: 4300  len: 49    finds: 1   ! Minigene HLA-DRalpha chain/
    GENESEQP:R49589  ck: 4300  len: 49    finds: 1   ! Sequence of HLA-DR alpha ch
    GENESEQP:R50081  ck: 3747  len: 352   finds: 1   ! NANBH virus antigenic fragm
    GENESEQP:R50082  ck: 108   len: 215   finds: 1   ! NANBH virus antigenic fragm
    GENESEQP:R53251  ck: 2035  len: 56    finds: 1   ! Consensus sequence of signa
    GENESEQP:R53646  ck: 4677  len: 380   finds: 1   ! c-fos gene product. New ant
    GENESEQP:R53748  ck: 2863  len: 355   finds: 1   ! Seven transmembrane recepto
    GENESEQP:R58556  ck: 4528  len: 175   finds: 1   ! Heparin-binding secretory t
    GENESEQP:R63592  ck: 2807  len: 79    finds: 1   ! MAP-kinase-phosphatase CL10
    GENESEQP:R63602  ck: 5661  len: 367   finds: 1   ! MAP-kinase-phosphatase CL10
    GENESEQP:R58635  ck: 2685  len: 124   finds: 1   ! Amylase inhibitor protein 0
    GENESEQP:R58636  ck: 3320  len: 124   finds: 1   ! Amylase inhibitor protein 0
    GENESEQP:R58587  ck: 9708  len: 246   finds: 1   ! Nicotineamide adenine dinuc
    GENESEQP:R63443  ck: 8722  len: 120   finds: 1   ! Amino terminal sequence of
    GENESEQP:R63444  ck: 7811  len: 120   finds: 1   ! Amino terminal sequence of
    GENESEQP:R62507  ck: 2303  len: 313   finds: 1   ! Galactosyl transferase 3' c
    GENESEQP:R66355  ck: 1046  len: 216   finds: 1   ! Ii protein. Identification
    GENESEQP:R63159  ck: 2293  len: 376   finds: 1   ! Mouse growth differentiatio
    GENESEQP:R63160  ck: 1814  len: 375   finds: 1   ! Human growth differentiatio
    GENESEQP:R66665  ck: 4995  len: 25    finds: 1   ! Binding peptide B (84 TS) I
    GENESEQP:R66680  ck: 5136  len: 25    finds: 1   ! Peptide 84 TS-CYS. Immunoge
    GENESEQP:R66682  ck: 5012  len: 25    finds: 1   ! Peptide MONOSER. Immunogeni
    GENESEQP:R67596  ck: 5465  len: 331   finds: 1   ! A. aculeatus pectin methyle
    GENESEQP:R71324  ck: 7564  len: 329   finds: 1   ! Acetyl-CoA-reductase. Trans
    GENESEQP:R70100  ck: 4652  len: 274   finds: 1   ! Lettuce infectious yellows
    GENESEQP:R72715  ck: 516   len: 359   finds: 1   ! hisC gene product of Salmon
    GENESEQP:R74180  ck: 8163  len: 248   finds: 1   ! Type I ribosome-inactivatin
```

GENESEQP:R84103	ck: 5723	len: 303	finds: 1	! Equine herpesvirus (EHV1) uniq
GENESEQP:R81323	ck: 9504	len: 123	finds: 1	! Humanized VLA-4 antibody 21.6
GENESEQP:R81327	ck: 2629	len: 140	finds: 1	! Mouse VLA-4 antibody 21.6 ligh
GENESEQP:R81330	ck: 8013	len: 123	finds: 1	! Mouse anti-VLA-4 antibody 21.6
GENESEQP:R81333	ck: 9019	len: 142	finds: 1	! Human VLA-4 reshaped antibody
GENESEQP:R75202	ck: 824	len: 225	finds: 1	! Tyrosine phosphatase MPTP-delt
GENESEQP:R91274	ck: 1456	len: 62	finds: 1	! Eimeria gametocyte antigen. Pr
GENESEQP:R94989	ck: 5431	len: 354	finds: 1	! Nsk2 receptor intracellular do
GENESEQP:R94990	ck: 5928	len: 282	finds: 1	! Nsk2 receptor tyrosine-kinase
GENESEQP:R89896	ck: 6122	len: 423	finds: 1	! Rat kynurenine aminotransferas
GENESEQP:R89897	ck: 206	len: 437	finds: 1	! Rat kynurenine aminotransferas
GENESEQP:R89898	ck: 7276	len: 457	finds: 1	! Rat kynurenine aminotransferas
GENESEQP:R77433	ck: 2940	len: 160	finds: 1	! Antigen of Rochalimaea hensela
GENESEQP:R95054	ck: 9286	len: 342	finds: 1	! TGF-a-DETA-DGAL4 multidomain p
GENESEQP:R95055	ck: 7433	len: 421	finds: 1	! IL-2-DETA-DGAL4 multidomain pr
GENESEQP:R91225	ck: 7811	len: 328	finds: 1	! Human placenta G-protein coupl
GENESEQP:R99695	ck: 4995	len: 25	finds: 1	! VP6 binding peptide B. Encapsu
GENESEQP:R99501	ck: 2030	len: 244	finds: 1	! Nitrilase regulatory factor 24
GENESEQP:R99576	ck: 2766	len: 83	finds: 1	! Wasp venom BrhTX-1 subunit (b)
GENESEQP:W00624	ck: 8391	len: 358	finds: 1	! SacI endonuclease. DNA encodin
GENESEQP:R94579	ck: 2400	len: 488	finds: 1	! Chlamydia pneumoniae polypepti
GENESEQP:R94580	ck: 3773	len: 271	finds: 1	! C. pneumoniae polypeptide anti
GENESEQP:R94586	ck: 9728	len: 259	finds: 1	! C. pneumoniae polypeptide anti
GENESEQP:R99316	ck: 4323	len: 458	finds: 1	! Human SH-PTP1 variant derived
GENESEQP:R83015	ck: 5439	len: 371	finds: 1	! Human thyroid transcription fa
GENESEQP:W00725	ck: 1679	len: 207	finds: 1	! Vascular endothelial growth fa
GENESEQP:R99253	ck: 4145	len: 374	finds: 1	! Cytoplasmic antiproteinase-2 p
GENESEQP:W01743	ck: 9728	len: 259	finds: 1	! C. pneumoniae 53 kDa antigen.
GENESEQP:W04831	ck: 1679	len: 207	finds: 1	! Vascular endothelial growth fa
GENESEQP:W11481	ck: 5195	len: 205	finds: 1	! D. immitis mature venom allerg
GENESEQP:W11479	ck: 9389	len: 221	finds: 1	! D. immitis venom allergen anti
GENESEQP:W13272	ck: 2018	len: 244	finds: 1	! Rhodococcus erythropolis SK92-
GENESEQP:W06943	ck: 9316	len: 140	finds: 1	! CagI locus product 14. Helicob
GENESEQP:W06935	ck: 8017	len: 136	finds: 1	! CagI locus product 6. Helicoba
GENESEQP:W20571	ck: 6963	len: 114	finds: 1	! H. pylori secreted or periplas
GENESEQP:W20886	ck: 2872	len: 154	finds: 1	! H. pylori secreted or periplas
GENESEQP:W20776	ck: 6185	len: 276	finds: 1	! H. pylori flagella-associated
GENESEQP:W20259	ck: 5141	len: 108	finds: 1	! H. pylori secreted or perip
GENESEQP:W20324	ck: 6372	len: 102	finds: 1	! H. pylori cell envelope tra
GENESEQP:W20696	ck: 7893	len: 121	finds: 1	! H. pylori secreted or perip
GENESEQP:W20139	ck: 1367	len: 53	finds: 1	! H. pylori cytoplasmic prote
GENESEQP:W09433	ck: 8653	len: 328	finds: 1	! Human placenta purinergic P
GENESEQP:W21704	ck: 7673	len: 248	finds: 1	! Luffin-A. Inactive precurso
GENESEQP:W21709	ck: 7681	len: 250	finds: 1	! Luffin-B. Inactive precurso
GENESEQP:W10031	ck: 8373	len: 74	finds: 1	! Protein encoded by clone G-
GENESEQP:W13647	ck: 5362	len: 30	finds: 1	! Invariant chain region whic
GENESEQP:W25141	ck: 7673	len: 248	finds: 1	! Luffin-A (a ribosome inhibi
GENESEQP:W25146	ck: 7681	len: 250	finds: 1	! Luffin-B (a ribosome inhibi
GENESEQP:W22410	ck: 2629	len: 140	finds: 1	! Alpha-4 integrin mouse MAb
GENESEQP:W22413	ck: 9557	len: 123	finds: 1	! Humanised alpha-4 integrin
GENESEQP:W22428	ck: 9019	len: 142	finds: 1	! Humanised alpha-4 integrin
GENESEQP:W26413	ck: 3737	len: 121	finds: 1	! Swinepox virus HindIII C en
GENESEQP:W17971	ck: 6033	len: 47	finds: 1	! RAC protein kinase C-termin
GENESEQP:W34668	ck: 9097	len: 438	finds: 1	! Arabidopsis thaliana Rec-A
GENESEQP:W24491	ck: 1423	len: 122	finds: 1	! Novel amylase inhibitor pro
GENESEQP:W24492	ck: 2685	len: 124	finds: 1	! Novel amylase inhibitor pro
GENESEQP:W24493	ck: 3320	len: 124	finds: 1	! Novel amylase inhibitor pro
GENESEQP:W29523	ck: 2685	len: 124	finds: 1	! Wheat amylase inhibitor 0.2
GENESEQP:W29524	ck: 1423	len: 122	finds: 1	! Wheat amylase inhibitor 0.2
GENESEQP:W29525	ck: 3320	len: 124	finds: 1	! Wheat amylase inhibitor 0.1
GENESEQP:R63906	ck: 7673	len: 248	finds: 1	! Type I ribosome-inactivatin
GENESEQP:W45093	ck: 9381	len: 176	finds: 1	! Residues 217-392 of human t
GENESEQP:W38190	ck: 1840	len: 107	finds: 1	! Soybean SCARECROW SRPg1 pro
GENESEQP:W34214	ck: 5054	len: 271	finds: 1	! Streptomyces ketoreductase
GENESEQP:W41747	ck: 3783	len: 239	finds: 1	! Hepatitis C virus antigen.
GENESEQP:W47606	ck: 9828	len: 194	finds: 1	! HRSV protein 22K. Productio
GENESEQP:W50006	ck: 5567	len: 365	finds: 1	! Human papillomavirus-18 E2
GENESEQP:W55590	ck: 428	len: 307	finds: 1	! H. pylori ORF 09cp10224_429
GENESEQP:W55815	ck: 5054	len: 271	finds: 1	! Streptomyces roseofulvus fr
GENESEQP:W56137	ck: 7408	len: 186	finds: 1	! Open reading frame 2 peptid
GENESEQP:W51011	ck: 7216	len: 277	finds: 1	! Human liver carbonyl reduct
GENESEQP:W28052	ck: 2273	len: 269	finds: 1	! Amino acid sequence of phos
GENESEQP:W37816	ck: 3813	len: 317	finds: 1	! Human secreted apoptosis-re

GENESEQP:W59024	ck: 9911	len: 71	finds: 1	! Enterocin-900. Bacterial growt
GENESEQP:W57330	ck: 4443	len: 339	finds: 1	! Glycerol-3-phosphate dehydroge
GENESEQP:W62641	ck: 7060	len: 137	finds: 1	! Flea serine protease inhibitor
GENESEQP:W55052	ck: 4306	len: 118	finds: 1	! Sunflower antifungal protein M
GENESEQP:W48722	ck: 2863	len: 355	finds: 1	! Human V28 seven transmembrane
GENESEQP:W54728	ck: 3126	len: 24	finds: 1	! Peptide from Ii p80-103. Incre
GENESEQP:W60258	ck: 4443	len: 339	finds: 1	! Klebsiella pneumoniae glycerol
GENESEQP:W60267	ck: 843	len: 387	finds: 1	! Klebsiella pneumoniae DHAT prc
GENESEQP:W60764	ck: 7422	len: 260	finds: 1	! Rainbow trout interleukin 1 be
GENESEQP:W57889	ck: 5290	len: 271	finds: 1	! Corn raffinose synthetase. New
GENESEQP:W44905	ck: 1135	len: 39	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44906	ck: 6449	len: 70	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44907	ck: 3740	len: 127	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44908	ck: 1205	len: 79	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44909	ck: 4699	len: 110	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44910	ck: 1990	len: 167	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44879	ck: 9655	len: 104	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W44882	ck: 8918	len: 61	finds: 1	! "polyproline beta-turn helix"
GENESEQP:W62612	ck: 8912	len: 369	finds: 1	! Human glutamate-binding protei
GENESEQP:W30679	ck: 843	len: 387	finds: 1	! 1,3-Propanediol oxidoreductase
GENESEQP:W69232	ck: 3029	len: 263	finds: 1	! FcR-II protein sequence. Nucle
GENESEQP:W62617	ck: 5610	len: 212	finds: 1	! Rattus norvegicus SOCS1 protei
GENESEQP:W62622	ck: 4000	len: 130	finds: 1	! Mus musculus SOCS8 protein. Su
GENESEQP:W62623	ck: 6307	len: 207	finds: 1	! Homo sapiens SOCS11 protein. S
GENESEQP:W62624	ck: 8778	len: 134	finds: 1	! Mus musculus SOCS13 protein. S
GENESEQP:W62613	ck: 4527	len: 212	finds: 1	! Mus musculus SOCS1 protein. Su
GENESEQP:W62614	ck: 1060	len: 198	finds: 1	! Mus musculus SOCS2 protein. Su
GENESEQP:W62615	ck: 9828	len: 225	finds: 1	! Mus musculus SOCS3 protein. Su
GENESEQP:W62616	ck: 758	len: 211	finds: 1	! Homo sapiens SOCS1 protein. Su
GENESEQP:W61619	ck: 1162	len: 250	finds: 1	! Clone HTPEF86 of TM4SF superfa
GENESEQP:W65460	ck: 1814	len: 375	finds: 1	! Human growth differentiation f
GENESEQP:W64373	ck: 3575	len: 120	finds: 1	! Mycobacterium tuberculosis ant
GENESEQP:W38580	ck: 5404	len: 174	finds: 1	! Streptococcus pneumoniae prote
GENESEQP:W70962	ck: 4556	len: 212	finds: 1	! A STAT function regulatory prc
GENESEQP:W77666	ck: 5825	len: 100	finds: 1	! Staphylococcus aureus protein
GENESEQP:W38716	ck: 3773	len: 161	finds: 1	! S. pneumoniae aspartyl tRNA sy
GENESEQP:W79302	ck: 8790	len: 67	finds: 1	! A Staphylococcus aureus protei

GENESEQP:W69889	ck: 1496	len: 376	finds: 1	! Rat growth differentiation
GENESEQP:W69891	ck: 1805	len: 375	finds: 1	! Pig growth differentiation
GENESEQP:W69892	ck: 1229	len: 375	finds: 1	! Ovine growth differentiatio
GENESEQP:W30689	ck: 2293	len: 376	finds: 1	! Murine growth differentiati
GENESEQP:W69885	ck: 1814	len: 375	finds: 1	! Human growth differentiatio
GENESEQP:W69886	ck: 1463	len: 375	finds: 1	! Baboon growth differentiati
GENESEQP:W69887	ck: 9305	len: 375	finds: 1	! Bovine growth differentiati
GENESEQP:W80799	ck: 2018	len: 244	finds: 1	! Rhodococcus nitrile hydrata
GENESEQP:W80624	ck: 4504	len: 189	finds: 1	! S. pneumoniae GTP binding p
GENESEQP:W68494	ck: 9982	len: 368	finds: 1	! E2 papillomavirus protein e
GENESEQP:W82504	ck: 447	len: 225	finds: 1	! Human EPRG1 protein #1. New
GENESEQP:W82505	ck: 9756	len: 157	finds: 1	! Human EPRG1 protein #2. New
GENESEQP:W81740	ck: 3575	len: 120	finds: 1	! M. tuberculosis immunogenic
GENESEQP:W72902	ck: 5027	len: 226	finds: 1	! Mycobacterium tuberculosis
GENESEQP:W80495	ck: 1679	len: 207	finds: 1	! Human vascular endothelial
GENESEQP:W85306	ck: 3126	len: 24	finds: 1	! Helper T-cell class II pept
GENESEQP:W86202	ck: 9273	len: 206	finds: 1	! Human VEGF-related factor (
GENESEQP:W68541	ck: 8848	len: 420	finds: 1	! Amino acid sequence of the
GENESEQP:W86235	ck: 8227	len: 185	finds: 1	! Human VRF (VEGF-related fac
GENESEQP:W86214	ck: 5797	len: 178	finds: 1	! Human VRF-2 truncated fragm
GENESEQP:W86215	ck: 4363	len: 173	finds: 1	! Human VRF-2 truncated fragm
GENESEQP:W86216	ck: 2203	len: 168	finds: 1	! Human VRF-2 truncated fragm
GENESEQP:W86217	ck: 3123	len: 163	finds: 1	! Human VRF-2 truncated fragm
GENESEQP:W85119	ck: 5663	len: 446	finds: 1	! A delta-5 desaturase enzyme
GENESEQP:W73507	ck: 3945	len: 317	finds: 1	! Human ATG-1709 protein. New
GENESEQP:W90011	ck: 9038	len: 343	finds: 1	! Expressed antigen for clust
GENESEQP:W89910	ck: 8829	len: 343	finds: 1	! Antigen 1 from cluster 35a.
GENESEQP:W85590	ck: 2306	len: 354	finds: 1	! Polypeptide with 98% homolo
GENESEQP:W85591	ck: 8607	len: 351	finds: 1	! Polypeptide with 95% homolo
GENESEQP:W89100	ck: 1342	len: 36	finds: 1	! Polypeptide fragment encode
GENESEQP:W88568	ck: 9823	len: 221	finds: 1	! Secreted protein encoded by
GENESEQP:W83391	ck: 9212	len: 161	finds: 1	! Caenorhabditis elegans synM
GENESEQP:W95506	ck: 5663	len: 446	finds: 1	! Mortierella alpina delta 5
GENESEQP:W98862	ck: 4789	len: 304	finds: 1	! H. pylori GHPO 1670 protein
GENESEQP:W98755	ck: 3203	len: 102	finds: 1	! H. pylori GHPO 1075 protein
GENESEQP:W98226	ck: 2117	len: 115	finds: 1	! H. pylori GHPO 1099 protein

```
    GENESEQP:W88330  ck: 7734  len: 336    finds: 1    ! Mannosyl transferase involved
    GENESEQP:Y01440  ck: 9731  len: 64     finds: 1    ! Secreted protein encoded by ge
    GENESEQP:W99185  ck: 9793  len: 314    finds: 1    ! Rhodococcus corallina ohp oper
    GENESEQP:W97722  ck: 1929  len: 302    finds: 1    ! Staphylococcus aureus mutant P
    GENESEQP:Y11009  ck: 2840  len: 142    finds: 1    ! H. pylori ORF hp6p22217_235640
    GENESEQP:Y11062  ck: 4555  len: 116    finds: 1    ! H. pylori ORF 06cp30603_107440
    GENESEQP:W97887  ck: 1814  len: 375    finds: 1    ! Human myostatin. Increasing mu
    GENESEQP:W97884  ck: 8371  len: 375    finds: 1    ! Bovine myostatin. Increasing m
    GENESEQP:W97885  ck: 1462  len: 286    finds: 1    ! Bovine myostatin (mutant form)
    GENESEQP:W97886  ck: 2420  len: 376    finds: 1    ! Murine myostatin. Increasing m
    GENESEQP:W97840  ck: 7697  len: 207    finds: 1    ! Winged bean chymotrypsin inhib
    GENESEQP:Y03794  ck: 7277  len: 144    finds: 1    ! S. aureus polypeptide. New ess
    GENESEQP:Y03782  ck: 881   len: 135    finds: 1    ! S. aureus polypeptide. New ess
    GENESEQP:Y02934  ck: 2639  len: 53     finds: 1    ! Fragment of human secreted pro
    GENESEQP:Y11970  ck: 1817  len: 113    finds: 1    ! Human 5' EST secreted protein
    GENESEQP:W96262  ck: 3324  len: 423    finds: 1    ! Brn-3a polypeptide. New polype
    GENESEQP:Y05204  ck: 3125  len: 24     finds: 1    ! Human CLIP immunomodulatory pe
    GENESEQP:Y11993  ck: 4119  len: 102    finds: 1    ! Human 5' EST secreted protein
    GENESEQP:Y05545  ck: 5919  len: 243    finds: 1    ! Wheat Type III glutathione tra
    GENESEQP:Y07041  ck: 1679  len: 207    finds: 1    ! Breast cancer associated antig
    GENESEQP:Y07007  ck: 4692  len: 177    finds: 1    ! Breast cancer associated antig
    GENESEQP:W93953  ck: 7759  len: 352    finds: 1    ! Human regulatory molecule HRM-
    GENESEQP:Y04959  ck: 8454  len: 273    finds: 1    ! Mycobacterium species protein
    GENESEQP:Y04960  ck: 8320  len: 280    finds: 1    ! Mycobacterium species protein
    GENESEQP:Y04901  ck: 6286  len: 72     finds: 1    ! Mycobacterium species protein
    GENESEQP:Y04743  ck: 5429  len: 110    finds: 1    ! Mycobacterium species protein
    GENESEQP:Y07239  ck: 2144  len: 141    finds: 1    ! Fragment of human placental ST
    GENESEQP:Y07240  ck: 2614  len: 141    finds: 1    ! Fragment of human hepatic STAT
    GENESEQP:Y07241  ck: 3064  len: 141    finds: 1    ! Fragment of mouse hepatic STAT
\\End of list

 Databases searched:
       GeneSeq, Release 36.3, Released on 8Jul1999, Formatted on 9Jul1999

    Total finds:        245
   Total length: 23,686,106
 Total sequences:    188,963
       CPU time:   03:32.95

! FINDPATTERNS on geneseqp:* allowing 0 mismatches

!       1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C,T,S)(R,K,H
```

```
              P90009  ck: 7203  len: 102    ! Mouse gamma-2b chain New immunoglobulin
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{6}(P)x(P)(A)xx(L)(S)(R)xx(V)xxx(C)x{80}
              1:        PQVYILPPPAEQLSRKDVSLTCLVVGFNPGDISVEWTSNGHTEENYKDTAPVLDSDGSYFI

              P91884  ck: 718   len: 48     ! Antigenic Epstein-Barr virus peptide Ide
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{19}(P)x(P)(L)xx(A)(S)(R)xx(P)xxx(S)x{13}
              1:        RGWFCPSLCPSEEPGTSGTPEPLGPASRRPPGLRSPLSPVKPKECLRG

              P94157  ck: 2829  len: 246    ! Acetyl-CoA reductase. Constructing new p
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{112}(L)x(S)(L)xx(V)(T)(K)xx(I)xxx(A)x{118}
              1:      MTQRIAYVTGGMGGIGTAICQRLAKDGFRVVAGCGPNSPRREKWLEQQKALGFDFIASEGNVA
VNTVSPGYIATDMVKAIRQDVLDKIVATIPVKRLGLPEEIASICAWLSSEESGFST
GADFSLNGGLHMG

              P90460  ck: 560   len: 288    ! Alpha-factor profibrolase from Southern
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{149}(L)x(T)(V)xx(V)(S)(H)xx(L)xxx(G)x{123}
              1:      MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVLPFSNSTNNGL
HDNAQLLTAIDFDGDTVGLAYVGGMCQLKHSTGVIQDHSAINLLVALTMAHELGHN
LGMNHDGNQCHCGANSCVMAAMLSDQPSKLFSDCSKKDYQTFLTVNNPQCILNKP

              P81999  ck: 4995  len: 25     ! Binding peptide 2 capable of binding to
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{2}(I)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{7}
              1:        CNIAPASIVSRNIVYTRAQPNQDIA

              P70477  ck: 9828  len: 194    ! Sequence of human respiratory syncytial
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{83}(I)x(S)(I)xx(I)(T)(K)xx(A)xxx(M)x{95}
              1:      MSRRNPCKFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDKSIDTLSEISGA
ESTVSDTNDHAKNNDTT

              P70785  ck: 9828  len: 194    ! Sequence encoding human respiratory sync
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{83}(I)x(S)(I)xx(I)(T)(K)xx(A)xxx(M)x{95}
              1:      MSRRNPCKFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDKSIDTLSEISGA
ESTVSDTNDHAKNNDTT

              R24296  ck: 3600  len: 384    ! Regulatory protein VanS involved in glyc
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{84}(V)x(S)(I)xx(L)(C)(R)xx(L)xxx(A)x{284}
              1:      LVIKLKNKKNDYSKLERKLYMYIVAIVVVAIVFVLYIRSMIRGKLGDWILSILENKYDLNHLD
SLLDEAPDMPVDQKAKYVHITLDKAYRLEQLIDEFFEITRYNLQTITLTKTHIDLY
YMLVQMTDEFYPQLSAHGKQAVIHAPEDLTVSGDPDKLARVFNNILKNAAAYSEDNSIIDITAGLSGDVVSIEFKNTGSIPKDKL

              R10995  ck: 1620  len: 398    ! Xenopus Bone Morphogenetic Factor BMP-2A
1                     <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                        x{178}(V)x(P)(A)xx(A)(S)(R)xx(V)xxx(L)x{204}
              1:      MVAGIHSLLLLQFYQILLSGCTGLVPEEGKRKYSESTRSSPQQSQQVLDQFELRLLNMFGLKR
IVKPAAAASRGPVVRLLDTRLIHHNESKWESFDVTPAITRWIAHKQPNHGFVVEVT
HLDNDTNVPKRHVRISRSLTLDKGHWPRIRPLLVTFSHDGKGHALHKRQKRQARHKQRKRLKSSCRRHPLYVDFSDVGWNDWIVA
```

R24190 ck: 5872 len: 186 ! Bovine RSV strain A 51908 M2 protein. Bovin

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{83}(L)x(S)(I)xx(I)(T)(K)xx(A)xxx(M)x{87}
1: MSRRNPCKYEIRGHCLNGKKCHFSHNYFEWAPHALLVRQNFMLNKILKSMDRNNDTLSEISGAAEL
DINVDEQNE

P81092 ck: 3060 len: 95 ! Sequence of rhinovirus HRV2 viral protein P

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{27}(I)x(P)(V)xx(I)(S)(K)xx(I)xxx(L)x{52}
1: GVTDYIHMLGEAFGNGFVDSVKEHIHAINPVGNISKKIIKWMLRIISAMVIIIRNSSDPQTILA

P81093 ck: 5281 len: 323 ! Sequence of rhinovirus HRV2 viral protein P

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{76}(I)x(T)(L)xx(L)(S)(R)xx(L)xxx(A)x{231}
1: SDSWLKKFTEACNAARGLEWIGNKISKFIEWMKSMLPQAQLKVKYLNELKKLNLYEKQVESLRAVD
GDDMTLFCQMVSSVTFIPPMADLPDKGKAFDSRFVLCSTNHSLLTPPTITSLPAMN
RRFFLDLDIIVHDNFKDPQGKLNVAAAFRPCDVDNRIGNARCCPFVCGKAVSFKDRNSCNKYSLAQVYNIMIEEDRRRRQVVDVMTAI

R10974 ck: 3836 len: 246 ! Acetoacetyl CoA reductase enzyme. Construct

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{112}(L)x(S)(L)xx(V)(T)(K)xx(I)xxx(A)x{118}
1: MTQRIAYVTGGMGGIGTAICQRLAKDGFRVVAGCGPNSPRREKWLEQQKALGYDFIASEGNVADWD
VNTVSPGYIATDMVKAIRQDVLDKIVATIPVKRLGLPEEIASICAWLSSEESGFST
GADFSLNGGLHMG

R12468 ck: 7673 len: 248 ! Luffa cylindrica bioactive protein. New prc

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
1: DVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNLTILLSSASGASRYTLMTLSNYDGKAITVAVDV
QVPSLATISLENSLWSALSKQIQLAQTNNGTFKTPVVITDDKGQRVEITNVTSKVV
TKNIQLLLNYKQNVA

R13252 ck: 528 len: 247 ! Murine Cytotoxic Cell Protease-1. DNA vecto

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{6}(L)x(T)(L)xx(A)(S)(R)xx(A)xxx(I)x{225}
1: MKILLLLLTLSLASRTKAGEIIGGHEVKPHSRPYMALLSIKDQQPEAICGGFLIREDFVLTAAHCE
KNRYNKTNQICAGDPKTKRASFRGDSGGPLVCKKVAAGIVSYGYKDGSPPRAFTKV
SSFLSWIKKTMKSS

P60645 ck: 2381 len: 292 ! Mouse kidney cell Band 3-like anion transpo

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{106}(A)x(T)(V)xx(V)(T)(H)xx(A)xxx(M)x{170}
1: VDYSIEDTYTQKLSVPSGFSVTAPDKRGWVINPLGEKTPFPVWMMVASLLPAVLVFILIFMETQIT
VTSLNGIQFYERLHLLLMPPKHQPDVTYVKKVRTMRMHLFTALQLLCLALLWAVMS
TAASLAFPFILILTVPLRMVVLTRIFTEREMKCLDANEAEPVFDECEGVDEYNEMPMPV

R14247 ck: 3078 len: 149 ! HCV-T (6110-6557) encoded epitope. New nucl

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{15}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{118}
1: VLKESKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSRAINHIRSVWKDLLEDTET

R20004 ck: 4995 len: 25 ! VP6 inner capsid epitope, 84 TS, Peptide B.

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{2}(I)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{7}
1: CNIAPASIVSRNIVYTRAQPNQDIA

R21409 ck: 617 len: 482 ! NADH dehydrogenase subunit 2. New DNA se

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{184}(L)x(S)(L)xx(F)(S)(K)xx(L)xxx(M)x{282}
1: MLLSSIISQLIAIALSSSHWNLVLLSRISIISLIYSIILTYNVYYVEIIGLGLGIYNGFLQVT
SYTGITSLESLAIFSKVNLNIYMQISLLICVLGILFKIGIVPFHQWAIDVYDGVPT
IITTWLTTLTKISLLIFLMEFIYHHSSENWTTILMLLSVLSVIVGSLLGLSQSRIKRLLIYSMVSHVGFLMLSLSIMTEKSLEAF

R23004 ck: 7059 len: 297 ! Protein transcribed from the mba sequenc

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{23}(A)x(S)(V)xx(I)(T)(R)xx(L)xxx(I)x{258}
1: MDFLINKKLKIFITLMETGSFSIATSVLYITRTPLSRVISDLERELKQRLFIRKNGTLIPTEF
GKPDINRLAGTPVLFHEGAKNFNLDTIYHFFEQTLGITNPAFSFDNVDLFSSLYRL
QQGLAMLLIPVRVCRALGLSTDHALHIKGVALCTSLYYPTKKRETPDYRKAIKLIQQELKQSTF

R22695 ck: 4919 len: 286 ! blaC31 mutation (5). Novel DNA modificat

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{193}(A)x(S)(V)xx(A)(S)(R)xx(L)xxx(M)x{77}
1: MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEER
TTMPAAMATTLRKLLTASSVTLASRQQLRDWMEADKVAGPLLRSALPAGWFIADKS
GAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW

R25601 ck: 5700 len: 85 ! NANBH virus antigen. Antigenic polypepti

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{32}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{37}
1: KKVTFDRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRN

R25869 ck: 3783 len: 239 ! HCV polypeptide 16. Hepatitis C virus an

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{84}(L)x(S)(I)xx(A)(C)(K)xx(P)xxx(A)x{139}
1: VCCSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHSMVYSTTSRSASLRQKKVTFDRLQVL
LYDVVSTLPQAVMGPSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTE
NDIRTE

R25891 ck: 4644 len: 253 ! HK16. Hepatitis C virus antigen expresse

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{96}(L)x(S)(I)xx(A)(C)(K)xx(P)xxx(A)x{141}
1: MITPSLHACMQVVCCSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHSMVYSTTSRSASLR
PDLGVRVCEKMALYDVVSTLPQAVMGPSYGFQYSPGQRVEFLVNTWKSKKCPMGFS
YDTRCFDSTVTENDIRTEGL

R25303 ck: 9668 len: 194 ! HRSV 22K protein. Vaccines for human res

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{83}(I)x(S)(I)xx(I)(T)(K)xx(A)xxx(I)x{95}
1: MSRTNPCKFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDKSIDTLSEISGA
ESTVSDTNDHAKNNDTT

R29874 ck: 9846 len: 171 ! HCV NS4-NS5 peptide N29-1, N29-2, N29-3

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
1: YRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIRSVWXDL

R29879 ck: 8945 len: 285 ! HCV NS4-NS5 peptide 2918. New hepatitis

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{251}
1: YRDVLKDMKAKASTVKAKLLSVEEACKLTPPHSARSKFDYGAKDVQSLSSKAVNHIHSVWKDL

QCCDLDPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCY
LKASAACRAAKLQDCTMLVCGDDLVVICESAGTQEDAANLRVFTEAMTRNSA

R29895 ck: 8862 len: 171 ! HCV NS4-NS5 peptide N29-1. New hepatitis C

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
1: YRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIRSVWKDLLED

R29896 ck: 8862 len: 171 ! HCV NS4-NS5 peptide N29-2. New hepatitis C

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{137}
1: YRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIRSVWKDLLED

R29897 ck: 2886 len: 167 ! HCV NS4-NS5 peptide N29-3. New hepatitis C

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{18}(L)x(S)(V)xx(A)(C)(K)xx(P)xxx(A)x{133}
1: YRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHIRSVWKDLLED

R29909 ck: 429 len: 277 ! Prod. of the luffin-f gene. New gene coding

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{57}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{204}
1: MKRFTVLILAIFVAASTVEADVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNITLLLSSASGASR
AEASRFKYIEGQIIERISKNQVPSLATISLENEWSALSKQIQLAQTNNGTFKTPVV
ITDDKGQRVEITNVTSKVVTKNIQLLLNYKQNVAAFDEDVSAKH

R29910 ck: 9157 len: 278 ! Prod. of the luffin-g gene. New gene coding

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{59}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{203}
1: MNRFTFLSLLIILIAFFTVEGANVSFSLSGADSKSYSKFITALRKALPSKEKVSNIPLLLPSASGA
TTAEASRFKYIEGQIIERIPKNEVPSPAALSLENEWSALSKQIQLAQTNNGAFRTP
VVIIDNKGQRVEIKDVNSKVVTNNIKLLLNKQNIAAFDDGIPTKH

R31608 ck: 7012 len: 205 ! Homologous to chicken nov gene exon 3-4-enc

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{110}(L)x(P)(A)xx(V)(S)(R)xx(L)xxx(V)x{79}
1: QIPTRIPDALDVRVPQCLTSASPTPLFPSSSPAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGA
LPGPDTEWSACSKTCGMGYSTRVTNDNA

R30618 ck: 7998 len: 106 ! Polypeptide coded by Korean HCV cDNA fragme

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{54}(P)x(P)(I)xx(I)(T)(K)xx(L)xxx(G)x{36}
1: LLTLSPYYKVFLARLIWWLQYFITRAEAHLQVWIPPLNVRGGRDAIILLACAVHPEPIFDITKY

R32192 ck: 2829 len: 246 ! Sequence encoded by the acetoacetyl-CoA red

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{112}(L)x(S)(L)xx(V)(T)(K)xx(I)xxx(A)x{118}
1: MTQRIAYVTGGMGGIGTAICQRLAKDGFRVVAGCGPNSPRREKWLEQQKALGFDFIASEGNVADWD
VNTVSPGYIATDMVKAIRQDVLDKIVATIPVKRLGLPEEIASICAWLSSEESGFST
GADFSLNGGLHMG

R37294 ck: 7673 len: 248 ! Plant type I RIP Luffin. Analogues of type

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
1: DVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNLTILLSSASGASRYTLMTLSNYDGKAITVAVDV
QVPSLATISLENSLWSALSKQIQLAQTNNGTFKTPVVITDDKGQRVEITNVTSKVV

TKNIQLLLNYKQNVA

R39259 ck: 8652 len: 391 ! Human somatostatin receptor-1. Somatosta

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{188}(I)x(P)(I)xx(F)(S)(R)xx(A)xxx(G)x{187}
1: MFPNGTASSPSSSPSPSPGSCGEGGGSRGPGAGAADGMEEPGRNASQNGTLSEGQGSAILISF
NLGVWVLSLLVILPIVVFSRTAANSDGTVACNMLMPEPAQRWLVGFVLYTFLMGFL
LPVGAICLCYVLIIAKMRMVALKAGWQQRKRSERKITLMVMMVVMVFVICWMPFYVVQLVNVFAEQDDATVSQLSVILGYANSCA

R39260 ck: 8110 len: 391 ! Murine somatostatin receptor-1. Somatost

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{188}(I)x(P)(I)xx(F)(S)(R)xx(A)xxx(G)x{187}
1: MFPNGTASSPSSSPSPSPGSCGEGACSRGPGSGAADGMEEPGRNASQNGTLSEGQGSAILISF
NLGVWVLSLLVILPIVVFSRTAANSDGTVACNMLMPEPAQRWLVGFVLYTFLMGFL
LPVGAICLCYVLIIAKMRMVALKAGWQQRKRSERKITLMVMMVVMVFVICWMPFYVVQLVNVFAEQDDATVSQLSVILGYANSCA

R42317 ck: 3489 len: 345 ! EBV VCA-p40 from ORF BdRF1. Peptide(s) a

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{51}(A)x(P)(L)xx(F)(S)(H)xx(I)xxx(P)x{278}
1: MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFSHQAI
KDIAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSY
EVPRYVPHPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAARNSGPREDTNPQQPTTEGHHRGKKLVQASASGVAQSKE

R52667 ck: 5723 len: 303 ! Equine herpesvirus US2 gene. New recombi

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{62}(L)x(T)(L)xx(A)(C)(R)xx(V)xxx(G)x{225}
1: MGVVLITVVTVVDRHKALPNSSIDVDGHLWEFLSRQCFVLASEPLGIPIVVRSADLYRFSSSL
EHNHPLKLRAAGGEDTVVGECGFSKHSSNSLVRPPTVKRVIYAVVDPARLREIPAP
GRPLPRRRPSEGGMRAPRRRSRAPAAARSTAAAATPPRPGDPRAPAARRAGDVTWMERLLWGVFGRTSTR

R49296 ck: 4901 len: 25 ! Invarient chain (Ii) position 97-121. No

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{8}
1: LPKPPKPVSKMRMATPLLMQALPMG

R49297 ck: 3126 len: 24 ! Invarient chain (Ii) position 97-120. No

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1: LPKPPKPVSKMRMATPLLMQALPM

R49299 ck: 1278 len: 23 ! Invarient chain (Ii) position 97-119. No

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{6}
1: LPKPPKPVSKMRMATPLLMQALP

R49438 ck: 4300 len: 49 ! Minigene HLA-DRalpha chain/Ii 24 residue

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{26}(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1: MAISGVPVLGFFIIAVLMSAQESWALPKPPKPVSKMRMATPLLMQALPM

R49589 ck: 4300 len: 49 ! Sequence of HLA-DR alpha chain leader pe

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{26}(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
1: MAISGVPVLGFFIIAVLMSAQESWALPKPPKPVSKMRMATPLLMQALPM

```
R50081  ck: 3747  len: 352   ! NANBH virus antigenic fragment #13. Nucleic
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{52}(L)x(T)(L)xx(A)(C)(R)xx(P)xxx(A)x{284}
             1:  LLRYHNKVYCTTSKSASLRAKKVTFDRMQVLDAHYDSVLKDIKLAASKVSARLLTLEEACRLTPPH
AWADKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPEEARTAIRSLTERLYV
GGPMFNSKGQACGYRRCRASGVLTTSMGNTITCYVKALAACKAAGIVAPTMLVCGDDLVVISESQGTEEDERNLRAFTEAMTRYSAPP

R50082  ck: 108   len: 215   ! NANBH virus antigenic fragment #14. Nucleic
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{126}(A)x(S)(L)xx(Y)(T)(H)xx(L)xxx(A)x{73}
             1:  RYSAPPGDPPRPEYDLELITSCSSNVSVAISPQGRRRYYLSRDPTTPIARAAWETVRHSPVNSWLG
GRYLFNWAVKTKLKLTPLPEARLLDLSSWFTVGAGGGD

R53251  ck: 2035  len: 56    ! Consensus sequence of signal peptide of sma
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{5}(L)x(S)(A)xx(A)(T)(R)xx(P)xxx(S)x{35}
             1:    MASSMLSSAAVATRTNPAQASMVAPFTGLKSAAFPVSRKQNLDITSIASNGGRVQC

R53646  ck: 4677  len: 380   ! c-fos gene product. New antitumour vaccines
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{99}(A)x(S)(A)xx(Y)(S)(R)xx(V)xxx(M)x{265}
             1:  MMFSGFNADYEASSSRCSSASPAGDSLSYYHSPADSFSSMGSPVNAQDFCTDLAVSSANFIPTVTA
ALQTEIANLLKEKEKLEFILAAHRPACKIPDDLGFPEEMSVASLDLTGGLPEVATP
ESEEAFTLPLLNDPEPKPSVEPVKSISSMELKTEPFDDFLFPASSRPSGSETARSVPDMDLSGSFYAADWEPLHSGSLGMGPMATELE

R53748  ck: 2863  len: 355   ! Seven transmembrane receptor (V28). DNA enc
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{80}(V)x(T)(L)xx(W)(T)(H)xx(I)xxx(G)x{259}
             1:  MDQFPESVTENFEYDDLAEACYIGDIVVFGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPKSVT
DYPEVLQEIWPVLRNVETNFLGFLLPLLIMSYCYFRIIQTLFSCKNHKKAKAIKLI
LLVVIVFFLFWTPYNVMIFLETLKLYDFFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAFAGEKFRRYLYHLYGKCLAVLCGRSVH

R58556  ck: 4528  len: 175   ! Heparin-binding secretory transforming fact
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{131}(V)x(P)(I)xx(V)(T)(H)xx(P)xxx(V)x{28}
             1:  MLAGEIAGVNWESGYLVGIKRQRRLYCNVGIGFHLQVLPDGRISGTHEENPYSLLEISTVERGVVS

R63592  ck: 2807  len: 79    ! MAP-kinase-phosphatase CL100 partial sequen
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{6}(L)x(S)(A)xx(A)(S)(R)xx(M)xxx(L)x{57}
             1:    ILPFLYLGSAYHASRKDMLDALGITALINVSANCPNHFEGHYQYKSIPVEDNHKADISSWFNEA

R63602  ck: 5661  len: 367   ! MAP-kinase-phosphatase CL100. Screening for
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{182}(L)x(S)(A)xx(A)(S)(R)xx(M)xxx(L)x{169}
             1:  MVMEVGTLDAGGLRALLGERAAQCLLLDCRSFFAFNAGHIAGSVNVRFSTIVRRRAKGAMGLEHIV
LPFLYLGSAYHASRKDMLDALGITALINVSANCPNHFEGHYQYKSIPVEDNHKADI
SSNFMEAIDFIDSIKNAGGRVFVHCQAGISRSATICLAYLMRTNRVKLDEAFEFVKQRRSIISPNFSFMGQLLQFESQVLPHSCSAEA

R58635  ck: 2685  len: 124   ! Amylase inhibitor protein 0.26 AI. New amyl
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{91}(A)x(S)(I)xx(V)(C)(K)xx(I)xxx(A)x{17}
             1:    SGPWMCYPGYAFKVPALPGCRPVLKLQCNGSQVPEAVLRDCCQQLADISEWCRCGALYSMLDSM

R58636  ck: 3320  len: 124   ! Amylase inhibitor protein 0.19 AI. New a
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{91}(A)x(S)(I)xx(V)(C)(R)xx(I)xxx(A)x{17}
             1:    SGPWMCYPGQAFQVPALPACRPLLRLQCNGSQVPEAVLRDCCQQLAHISEWCRCGALYSML

R58587  ck: 9708  len: 246   ! Nicotineamide adenine dinucleotide synth
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{158}(L)x(P)(L)xx(F)(T)(K)xx(V)xxx(G)x{72}
             1:  MQEKIDKLVQWLRDQVSSAGLNGAVVGISGGIDSAVVAHLIKRAFPDDSLGLIMPCKSNPKDM
GVPEEIIKKAPSAGLWEGQTDESEMGTTYEMIDKYLKGEEIPERDRKIIERLHERS
HHKRQLAIAPPKF

R63443  ck: 8722  len: 120   ! Amino terminal sequence of the rubisco a
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{4}(V)x(T)(V)xx(A)(T)(R)xx(L)xxx(G)x{100}
             1:    MATAVSTVGAATRAPLNLNGSSAGASVPTSGFLGSSLKKHTNVRFPSSSRTTSMTVKAAEN

R63444  ck: 7811  len: 120   ! Amino terminal sequence of the rubisco a
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{4}(V)x(T)(V)xx(A)(T)(R)xx(L)xxx(G)x{100}
             1:    MATAVSTVGAATRAPLNLNGSSAGASVPTSGFLGSSLKKHTNVRFPSSSRTTSMTVKAAEN

R62507  ck: 2303  len: 313   ! Galactosyl transferase 3' clone product.
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{28}(I)x(T)(L)xx(V)(T)(R)xx(P)xxx(T)x{269}
             1:  VPSSNSASQSPQAMTDPCSPRLSYLSKAILTLCFVTRKPPEVVTITRWKAPVVWEGTYNRAVL
NFGVETLGQSVAQLQAWWYKAHPDEFTYERPKESAAYIPFRQGDFYYHAAILGGTP
TQVLNITQECFKGILQDKENDIEAEWHDESGLNKYFLLNKPTKILSPEYCWDYHIGMSVDIRIVKGAWQKKEYNLVRNNI

R66355  ck: 1046  len: 216   ! Ii protein. Identification of mutant Ii
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{81}(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{119}
             1:  MDDQRDLISNNEQLPMLGRRPGAPESKCSRGALYTGFSILVTLLLAGQATTAYFLYQQQGRLD
SRHSLEQKPTDAPPKESLELEDPSSGLGVTKQDLGPVPM

R63159  ck: 2293  len: 376   ! Mouse growth differentiation factor-8 pr
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{66}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
             1:  MMQKLQMYVYIYLFMLIAAGPVDLNEGSEREENVEKEGLCNACAWRQNTRYSRIEAIKIQILS
IKPMKDGTRYTGIRSLKLDMSPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDE
NGHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVF

R63160  ck: 1814  len: 375   ! Human growth differentiation factor-8 pr
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
             1:  MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFL

R66665  ck: 4995  len: 25    ! Binding peptide B (84 TS) Immunogenic co
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{2}(I)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{7}
             1:    CNIAPASIVSRNIVYTRAQPNQDIA

R66680  ck: 5136  len: 25    ! Peptide 84 TS-CYS. Immunogenic complex c
```

```
1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x(I)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{8}
                1:     NISPASIVSRNIVYTRAQPNQSIAC


          R66682  ck: 5012  len: 25    ! Peptide MONOSER. Immunogenic complex contg.

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{2}(I)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{7}
                1:     CNIAPASIVSRNIVYTSAQPNQDIA


          R67596  ck: 5465  len: 331   ! A. aculeatus pectin methylesterase. Enzyme

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{11}(A)x(S)(A)xx(A)(S)(R)xx(A)xxx(A)x{304}
                1:     MVKSVLASATFAVSATAASRTTAPSGAIVVAKSGGDYTTIGDAIDALSTSTTDTQTIFIEEGTYDE
IEGAVDFIFGQHARAWFQNVDIRVVEGPTSASITANGRSSETDTSYYVINKSTVAA
KEGDDVAEGTYYTGRPWSEYARVVFQQTSMTNVINSTGWTEWSTSTPNTEYVTFGEYANQPAPAPRAPAHVAEKTDAKTTITDITGSD

          R71324  ck: 7564  len: 329   ! Acetyl-CoA-reductase. Transgenic plant mate

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{195}(L)x(S)(L)xx(V)(T)(K)xx(I)xxx(A)x{118}
                1:     MASMISSSAVTTVSRASRGQSAAVAPFGGLKSMTGFPVKKVNTDITSITSNGGRVKCMQVWPPIGK
VVFRKMTRADWDAVIDTNLTSLFNVTKQVIDGMADRGWGRIVNISSVNGQKGQFGQ
TNYSTAKAGLHGFTMALAQEVATKGVTVNTVSPGYIATDMVKAIRQDVLDKIVATIPVKRLGLPEEIASICAWLSSEESGFSTGADFS

          R70100  ck: 4652  len: 274   ! Lettuce infectious yellows virus open readi

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{3}(M)x(P)(L)xx(L)(T)(K)xx(V)xxx(A)x{255}
                1:     MIMMSPLYALTKQCVIDTAYRLAVPTQHCAIYTVACRILFLSVGFMTIVKLCGFKMDTSSFIASIE
KVVTDLYLDLIGVLYVNTGYNIDLVEKFIFDKLEFLVYDGEEGFKSPQVEYNDICT
VYNLKPIIKYNRWHTDGSIVIECGDVIGKGINKTKKKICNK

          R72715  ck: 516   len: 359   ! hisC gene product of Salmonella typhimurium

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{170}(L)x(T)(L)xx(L)(T)(R)xx(A)xxx(A)x{173}
                1:     MSTENTLSVADLARENVRNLVPYQSARRLGGNGDVWLNANEFPTAVEFQLTQQTLNRYPECQPKAV
TRGKAIVVADEAYIEFCPQATLTGWLVEYPHLVILRTLSKAFALAGLRCGFTLANE
EVINLLLKVIAPYPLSTPVADIAAQALCPQGINAMRDRVAQTVQERQYLVNALQQTACVEHVFDSETNYILARFTASSSVFKSLWDQG

          R74180  ck: 8163  len: 248   ! Type I ribosome-inactivating protein luffin

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
                1:     DVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNLTILLSSASGASRYTLMTLSNYDGKAITVAVDV
QVPSLATISLENSLWSALSKQIQLAQTNNGTFKTPVVITDDKQQRVEITNVTSKVV
TKNIQLLLNYKQNVA

          R84103  ck: 5723  len: 303   ! Equine herpesvirus (EHV1) unique short (US2

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{62}(L)x(T)(L)xx(A)(C)(R)xx(V)xxx(G)x{225}
                1:     MGVVLITVVTVVDRHKALPNSSIDVDGHLWEFLSRQCFVLASEPLGIPIVVRSADLYRFSSSLLTL
EHNHPLKLRAAGGEDTVVGECGFSKHSSNSLVRPPTVKRVIYAVVDPARLREIPAP
GRPLPRRRPSEGGMRAPRRRSRAPAAARSTAAAATPPRPGDPRAPAARRAGDVTWMERLLWGVFGRTSTR

          R81323  ck: 9504  len: 123   ! Humanized VLA-4 antibody 21.6 heavy chain v

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{50}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{57}
                1:     QVQLVQSGAEVKKPGASVKVSCKASGFNIKDTYIHWVRQAPGQRLEWMGRIDPANGYTKYDPKF

          R81327  ck: 2629  len: 140   ! Mouse VLA-4 antibody 21.6 light heavy va

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{69}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{55}
                1:     MKCSWVMFFLMAVVTGVNSEVQLQQSGAELVKPGASVKLSCTASGFNIKDTYIHCVKQRPEQG

          R81330  ck: 8013  len: 123   ! Mouse anti-VLA-4 antibody 21.6 heavy cha

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{50}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{57}
                1:     EVQLQQSGAELVKPGASVKLSCTASGFNIKDTYIHCVKQRPEQGLEWIGRIDPANGYTKYD

          R81333  ck: 9019  len: 142   ! Human VLA-4 reshaped antibody 21.6 light

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{69}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{57}
                1:     MDWTWRVFCLLAVAPGAHSQVQLVQSGAEVKKPGASVKVSCKASGFNIKDTYIHWVRQAPGQR

          R75202  ck: 824   len: 225   ! Tyrosine phosphatase MPTP-delta insertio

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{26}(A)x(T)(P)xx(F)(T)(R)xx(V)xxx(G)x{183}
                1:     MPLTNCRMVPVRRPLSLLLTFFLCACAETPPRFTRTPVDQTGVSGGVASFICQATGDPRPKIV
NNGRIKQLRSESIGALQIEQSEESDQGKYECVATNSAGTRYSAPANLY

          R91274  ck: 1456  len: 62    ! Eimeria gametocyte antigen. Protection o

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{43}(P)x(P)(I)xx(A)(S)(K)xx(P)xxx(M)x{3}
                1:     LQHHHQHIILQQQQHHHHPPPAAVLVAAPPADKEQFAAAAAAAPIPIAAASKETPLIDMMS

          R94989  ck: 5431  len: 354   ! Nsk2 receptor intracellular domain (aa51

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{156}(P)x(T)(V)xx(L)(S)(H)xx(L)xxx(A)x{182}
                1:     YCCRRRKEWKNKKRESTAVTLTTLPSELLLDRLHPNPMYQRMPLLLNPKLLSLEYPRNNIEYV
GPPPLSCAEQLCIARQVAAGMAYLSERKFVHRDLATRNCLVGETMVVKIADFGLSR
NIYSADYYKADGNDAIPIRWMPPESIFYNRYTTESDVWAYGVVLWEIFSYGLQPYYGMAHEEVIYYVRDGNILACPENCPLELYN

          R94990  ck: 5928  len: 282   ! Nsk2 receptor tyrosine-kinase region (aa

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{97}(P)x(T)(V)xx(L)(S)(H)xx(L)xxx(A)x{169}
                1:     IEYVRDIGEGAFGRVFQARAPGLLPYEPFTMVAVKMLKEEASADMQADFQREAALMAEFDNPN
SADYYKADGNDAIPIRWMPPESIFYNRYTTESDVWAYGVVLWEIFSYGLQPYYGMA
HEEVIYYVRDGNILACPENCPLELYNLMRLCWSKLPADRPSFCSIHRIL

          R89896  ck: 6122  len: 423   ! Rat kynurenine aminotransferase (KAT) cl

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{161}(L)x(P)(A)xx(A)(S)(K)xx(P)xxx(V)x{246}
                1:     MTKRLQARRLDGIDQNLWVEFGKLTKEYDVVNLGQGFPDFSPPDFATQAFQQATSGNFMLNQY
LVLNTPNNPLGKVFSRMELELVANLCQQHDVVCISDEVYQWLVYDGHQHVSIASLP
GMWDRTLTIGSAGKSFSATGWKVGWVMGPDNIMKHLRTVHQNSIFHCPTQAQAAVAQCFEREQQHFGQPSSYFLQLPQAMELNRD

          R89897  ck: 206   len: 437   ! Rat kynurenine aminotransferase (KAT) cl

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{175}(L)x(P)(A)xx(A)(S)(K)xx(P)xxx(V)x{246}
                1:     MNHSTWCLFFKKLTMTKRLQARRLDGIDQNLWVEFGKLTKEYDVVNLGQGFPDFSPPDFATQA
PAELASKFTPRTKVLVLNTPNNPLGKVFSRMELELVANLCQQHDVVCISDEVYQWL
VYDGHQHVSIASLPGMWDRTLTIGSAGKSFSATGWKVGWVMGPDNIMKHLRTVHQNSIFHCPTQAQAAVAQCFEREQQHFGQPSS
```

```
R89898  ck: 7276  len: 457   ! Rat kynurenine aminotransferase (KAT) clone
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{195}(L)x(P)(A)xx(A)(S)(K)xx(P)xxx(V)x{246}
             1:       MFRSAAALSVHLMWPLWGRKAGASLTRCLHQSLTMTKRLQARRLDGIDQNLWVEFGKLTKEYDVVN
LKPSPAPKGKLGASNDWQLDPAELASKFTPRTKVLVLNTPNNPLGKVFSRMELELV
ANLCQQHDVVCISDEVYQWLVYDGHQHVSIASLPGMWDRTLTIGSAGKSFSATGWKVGWVMGPDNIMKHLRTVHQNSIFHCPTQAQAA

R77433  ck: 2940  len: 160   ! Antigen of Rochalimaea henselae. Nucleic ac
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{20}(L)x(S)(L)xx(I)(S)(H)xx(A)xxx(T)x{124}
             1:       MAAYISSKERKSMKKYSLVTLLSLFCISHAKAQTATLTDEYYKKALENTQKLDVAKSQTAESIYES

R95054  ck: 9286  len: 342   ! TGF-a-DETA-DGAL4 multidomain protein. Nucle
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{78}(A)x(T)(A)xx(A)(C)(H)xx(L)xxx(T)x{248}
             1:       DYKDDDDKLGTGVVSHFNDCPDSHTQFCFHGTCRFLVQEDKPACVCHSGYVGARCEHADLLASLEH
NDEAGAANADEKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKR
SPLTRAHLTEVESRLERLEQLFLLIFPREDLDMILKMDSLQDIKALLTGLFVQDNVNKDAVTDRLASVETDMPLTLRQHRISATSSSE

R95055  ck: 7433  len: 421   ! IL-2-DETA-DGAL4 multidomain protein. Nuclei
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{157}(A)x(T)(A)xx(A)(C)(H)xx(L)xxx(T)x{248}
             1:       DYKDDDDKLHHHHHHKLAPTSSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKK
PRGWEQLEQCGYPVQRLVALYLAARLSWNQVDQVIRNALASPGSGGDLGEAIREQP
EQARLALTLAAAESERFVRQGTGNDEAGAANADEKLLSSIEQACDICRLKKLKCSKEKPKCAKCLKNNWECRYSPKTKRSPLTRAHLT

R91225  ck: 7811  len: 328   ! Human placenta G-protein coupled receptor p
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{179}(L)x(P)(P)xx(A)(T)(H)xx(P)xxx(A)x{133}
             1:       MEWDNGTGQALGLPPTTCVYRENFKQLLLPPVYSAVLAAGLPLNICVITQICTSRRALTRTAVYTL
YDLSPPALATHYMPYGMALTVIGFLLPFAALLACYCLLACRLCRQDGPAEPVAQER
RGKAARMAVVVAAAFAISFLPFHITKTAYLAVGSTPGVPCTVLEAFAAAYKGTRPFASANSVLDPILFYFTQKKFRRRPHELLQKLTA

R99695  ck: 4995  len: 25    ! VP6 binding peptide B. Encapsulated bioacti
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{2}(I)x(P)(A)xx(V)(S)(R)xx(V)xxx(A)x{7}
             1:       CNIAPASIVSRNIVYTRAQPNQDIA

R99501  ck: 2030  len: 244   ! Nitrilase regulatory factor 244-amino acid
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{179}(A)x(P)(M)xx(F)(S)(R)xx(L)xxx(V)x{49}
             1:       MAGADVHAQGGTNRRARILVVDDEKHVRTMVTWQLESENFDVVAAADGDAALRQVTESGPDLMVLL
MAASPMQVFSRRRLLLEVWRSSPDWQQDATVTEHVHRIRRKIEEDPTKPTILQTVR
GAGYRFDGERA

R99576  ck: 2766  len: 83    ! Wasp venom BrhTX-1 subunit (b) partial sequ
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{2}(L)x(T)(M)xx(I)(C)(K)xx(L)xxx(L)x{65}
             1:       FFLITMSIICKIILLVLLSWTSMVSSTLFTDRKWCGRADKTFGPSRSLGGGVGDCCRSHDSCGR

W00624  ck: 8391  len: 358   ! SacI endonuclease. DNA encoding SacI restri
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{76}(A)x(S)(L)xx(L)(C)(H)xx(V)xxx(S)x{266}
             1:       MGITIKKSTAEQVLRKAYEAAASDDVFLEDWIFLATSLREVDAPRTYTAALVTALLARACDDRVDF
AIVEASLIAETQSFVVSGHDVPRKLQACVAAGLDMVYSEVVSRRINDPSRDFPGDV
QVILDGDPLLTVEVRGKSVSWEGLEQFVSSATYAGFRRVALMVDAASHVSLMSADDLTSALERKYECIVKVNESVSSFLRDVFVWSPH

R94579  ck: 2400  len: 488   ! Chlamydia pneumoniae polypeptide antigen
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{189}(I)x(T)(L)xx(A)(T)(K)xx(L)xxx(A)x{283}
             1:       MSISSSSGPDNQKNIMSQVLTSTPQGVPQQDKLSGNETKQIQQTRQGKNTEMESDATIAGASG
EVIEIGLALAKAIQTLGEATKSALSNYASTQAQADQTNKLGLEKQAIKIDKEREEY
QEMKAAEQKSKDLEGTMDTVNTVMIAVSVAITVISIVAAIFTCGAGLAGLAARAAVRAAAAGGAAGAAAATTVATQITVQAVVQA

R94580  ck: 3773  len: 271   ! C. pneumoniae polypeptide antigen (polyp
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{189}(I)x(T)(L)xx(A)(T)(K)xx(L)xxx(A)x{66}
             1:       MSISSSSGPDNQKNIMSQVLTSTPQGVPQQDKLSGNETKQIQQTRQGKNTEMESDATIAGASG
EVIEIGLALAKAIQTLGEATKSALSNYASTQAQADQTNKLGLEKQAIKIDKEREEY
QEMKAAEQKSKDLEGTMDTVNTVMIAKGFELPWGPLIN

R94586  ck: 9728  len: 259   ! C. pneumoniae polypeptide antigen (polyp
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{189}(I)x(T)(L)xx(A)(T)(K)xx(L)xxx(A)x{54}
             1:       MSISSSSGPDNQKNIMSQVLTSTPQGVPQQDKLSGNETKQIQQTRQGKNTEMESDATIAGASG
EVIEIGLALAKAIQTLGEATKSALSNYASTQAQADQTNKLGLEKQAIKIDKEREEY
QEMKAAEQKSKDLEGTMDTVNTVMIA

R99316  ck: 4323  len: 458   ! Human SH-PTP1 variant derived from eryth
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{157}(P)x(S)(P)xx(V)(T)(H)xx(V)xxx(G)x{285}
             1:       MVRWFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQGDFSLSVRVGDQVTHIRIQNSGDFYDL
VGGLETFDSLTDLVEHFKKTGIEEASGAFVYLRQPYYATRVNAADIENRVLELNKK
QESEDTAKAGFWEEFESLQKQEVKNLHQRLEGQRPENKGKNRYKNILPFDHSRVILQGRDSNIPGSDYINANYIKNQLLGPDENA

R83015  ck: 5439  len: 371   ! Human thyroid transcription factor-1. Ol
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{142}(L)x(S)(L)xx(V)(S)(K)xx(A)xxx(S)x{213}
             1:       MSMSPKHTTPFSVSDILSPLEESYKKVGMEGGGLGAPLAAYRQGQAAPPTAAMQQHAVGHHGA
RRFKQQKYLSAPEREHLASMIHLTPTQVKIWFQNHRYKMKRQAKDKAAQQQLQQDS
GGGGGGGGTGCPQQQQAQQQSPRRVAVPVLVKDGKPCQAGAPAPGAASLQGHAQQQAQHQAQAAQAAAAAISVGSGGAGLGAHPG

W00725  ck: 1679  len: 207   ! Vascular endothelial growth factor-like
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{158}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:       MSPLLRRLLLAALLQLAPAQAPVSQPDAPGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGT
AAPSTTSALTPGPAAAAADAAASSVAKGGA

R99253  ck: 4145  len: 374   ! Cytoplasmic antiproteinase-2 protein. Hu
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{149}(A)x(T)(V)xx(L)(T)(K)xx(L)xxx(I)x{209}
             1:       MDDLCEANGTFAISLFKILGEEDNSRNVFFSPMSISSALAMVFMGAKGSTAAQMSQALCLYKD
KYTRGMLFKTNEEKKTVQMMFKEAKFKMGYADEVHTQVLELPYVEEELSMVILLPD
DNTDLAVVEKALTYEKFKAWTNSEKLTKSKVQVFLPRLKLEESYDLEPFLRRLGMIDAFDEAKADFSGMSTEKNVPLSKVAHKCF

W01743  ck: 9728  len: 259   ! C. pneumoniae 53 kDa antigen. Detection
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{189}(I)x(T)(L)xx(A)(T)(K)xx(L)xxx(A)x{54}
             1:       MSISSSSGPDNQKNIMSQVLTSTPQGVPQQDKLSGNETKQIQQTRQGKNTEMESDATIAGASG
EVIEIGLALAKAIQTLGEATKSALSNYASTQAQADQTNKLGLEKQAIKIDKEREEY
QEMKAAEQKSKDLEGTMDTVNTVMIA
```

```
            W04831  ck: 1679  len: 207   ! Vascular endothelial growth factor-B186. Va

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{158}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
                1:    MSPLLRRLLLAALLQLAPAQAPVSQPDAPGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGTVAK
AAPSTTSALTPGPAAAAADAAASSVAKGGA


            W11481  ck: 5195  len: 205   ! D. immitis mature venom allergen antigen 5-

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{157}(A)x(T)(L)xx(I)(C)(H)xx(P)xxx(M)x{32}
                1:    YECEGGKLTPTERKNIVTQINKYRSRLIRGKLKNKDGYLMPKGKNMLRMRWDCKLEKSAQNWANMC
IYELGNPCKHNKDCRTKRCSAKSGLCKK


            W11479  ck: 9389  len: 221   ! D. immitis venom allergen antigen 5-like pr

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{173}(A)x(T)(L)xx(I)(C)(H)xx(P)xxx(M)x{32}
                1:    MILLVIFPAIIVVVASYECEGGKLTPTERKNIVTQINKYRSRLIRGKLKNKDGYLMPKGKNMLRMR
VFICHYFPGGNMVKDLIYELGNPCKHNKDCRTKRCSAKSGLCKK


            W13272  ck: 2018  len: 244   ! Rhodococcus erythropolis SK92-B1 regulatory

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{179}(A)x(P)(M)xx(F)(S)(R)xx(L)xxx(V)x{49}
                1:    MAGADVHAQGGTNRRARILVVDDEKHVRTMVTWQLESENFDVVAAADGDAALRQVTESAPDLMVLD
MAASPMQVFSRRRLLLEVWRSSPDWQQDATVTEHVHRIRRKIEEDPTKPTILQTVR
GAGYRFDGERA


            W06943  ck: 9316  len: 140   ! CagI locus product 14. Helicobacter pylori

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{40}(L)x(S)(V)xx(V)(S)(R)xx(L)xxx(I)x{84}
                1:    KRIFIKLNYYLLGVLSLACLQSVNADQNTDIKDISPEDMALNSVGLVSRDQLKIEIPKETLEQKVA


            W06935  ck: 8017  len: 136   ! CagI locus product 6. Helicobacter pylori C

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{102}(L)x(S)(L)xx(I)(S)(K)xx(L)xxx(A)x{18}
                1:    NIYYFMLLYKFTALNYFFKNQNGTNQISKLKQNFLQFKYSFNKHLDKYSLYYRLFNISSIVIGFLI


            W20571  ck: 6963  len: 114   ! H. pylori secreted or periplasmic protein 8

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{80}(L)x(S)(L)xx(I)(S)(K)xx(L)xxx(A)x{18}
                1:      MKRPISKLKQNFLQFKHSFNKHLDKYSLYYRLFNISSIVIGFLIALFSYGAGVILVYPILFLFA


            W20886  ck: 2872  len: 154   ! H. pylori secreted or periplasmic protein,

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{54}(L)x(S)(V)xx(V)(S)(R)xx(L)xxx(I)x{84}
                1:    NNAKNKNKGVKSMKTNFYKIKLLFAWCLIIGMFNAPLNADQNTDIKDISPEDMALNSVGLVSRDQL


            W20776  ck: 6185  len: 276   ! H. pylori flagella-associated protein, 07ge

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{45}(I)x(S)(I)xx(I)(S)(K)xx(V)xxx(G)x{215}
                1:    RLKMATKLTPKQKAQLDELSMSEKIAILLIQVGEDTTGEILRHLDIDSITEISKQIVQLNGTDKQI
NLGEISPQVVKRVSTVLENKLESLTSYKIEVGGLRAVAEIFNRLGQKSAKTTLARI
ESVDNKLAGAIKEMMFTFEDIAKLDNFAIMRDFKSGGLKKTGL


            W20259  ck: 5141  len: 108   ! H. pylori secreted or periplasmic protein 2

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{42}(L)x(S)(V)xx(V)(S)(R)xx(L)xxx(I)x{50}
                1:    MKTNFYKIKLLFAWCLIIGMFNAPLNADQNTDIKDISPEDMALNSVGLVSRDQLKIEIPKE


            W20324  ck: 6372  len: 102   ! H. pylori cell envelope transporter prot

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{2}(A)x(S)(L)xx(V)(S)(K)xx(L)xxx(L)x{84}
                1:    MMAHSLILVSKTSLSNLLIFVVQPDGKLSMTDAAIDPNMTNSGLRWYRVNEIAEKFKLIKD


            W20696  ck: 7893  len: 121   ! H. pylori secreted or periplasmic protei

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{87}(L)x(S)(L)xx(I)(S)(K)xx(L)xxx(A)x{18}
                1:    TSSFKTKMKRPISKLKQNFLQFKHSFNKHLDKYSLYYRLFNISSIVIGFLIALFSYGAGVI


            W20139  ck: 1367  len: 53    ! H. pylori cytoplasmic protein, 14257751.

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{33}(L)x(T)(I)xx(L)(S)(H)xx(V)xxx(P)x{4}
                1:    MLDFDLVLFGATGDLAMRKLFVSLYEIYISFMVLKTILGLSHRGVRSYPMKSF


            W09433  ck: 8653  len: 328   ! Human placenta purinergic P-2u receptor,

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{179}(L)x(P)(P)xx(A)(T)(H)xx(P)xxx(A)x{133}
                1:    MEWDNGTDQALGLPPTTCVYRENFKQLLLPPVYSAVLAPALPLNICVITQICTSRRALTRTAV
YDLSPPALATHYMPYGMALTVIGFLLPFAALLACYCLLACRLCRQDGPAEPVAQER
RGKAARMAVVVAAVFGISFLPFHITKTAYLAVRSTPGVPCTVLEAFAAAYKGTRPFASANSVLDPILFYFTQKKFRRRPHELLQK


            W21704  ck: 7673  len: 248   ! Luffin-A. Inactive precursor of maize ri

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
                1:    DVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNLTILLSSASGASRYTLMTLSNYDGKAITVA
QVPSLATISLENSLWSALSKQIQLAQTNNGTFKTPVVITDDKGQRVEITNVTSKVV
TKNIQLLLNYKQNVA


            W21709  ck: 7681  len: 250   ! Luffin-B. Inactive precursor of maize ri

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{38}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{196}
                1:    ANVSFSLSGADSKSYSKFITALRKALPSKEKVSNIPLLLPSASGASRYILMQLSNYDAKAITM
NEVPSPAALSLENEAWSLLSKQIQLAQTNNGAFRTPVVIIDNKGQRVEITNLASKV
QIKDVNSKLLLNKQNIA


            W10031  ck: 8373  len: 74    ! Protein encoded by clone G-4-5 from acti

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{8}(V)x(P)(P)xx(A)(S)(R)xx(P)xxx(P)x{50}
                1:    WILRRLPHVHPPQGASREEPGRLPQGRRRRRQTVLLASIYLASLLPIEIVCMCGVCMHLPT


            W13647  ck: 5362  len: 30    ! Invariant chain region which contains th

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{13}
                1:    LPKPPKPVSKMRMATPLLMQALPMGALPQG


            W25141  ck: 7673  len: 248   ! Luffin-A (a ribosome inhibitory protein)

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
                1:    DVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNLTILLSSASGASRYTLMTLSNYDGKAITVA
QVPSLATISLENSLWSALSKQIQLAQTNNGTFKTPVVITDDKGQRVEITNVTSKVV
TKNIQLLLNYKQNVA
```

W25146 ck: 7681 len: 250 ! Luffin-B (a ribosome inhibitory protein) in

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{38}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{196}
1: ANVSFSLSGADSKSYSKFITALRKALPSKEKVSNIPLLLPSASGASRYILMQLSNYDAKAITMAID
NEVPSPAALSLENEAWSLLSKQIQLAQTNNGAFRTPVVIIDNKGQRVEITNLASKV
QIKDVNSKLLLNKQNIA

W22410 ck: 2629 len: 140 ! Alpha-4 integrin mouse MAb 21.6 VH region.

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{69}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{55}
1: MKCSWVMFFLMAVVTGVNSEVQLQQSGAELVKPGASVKLSCTASGFNIKDTYIHCVKQRPEQGLEW

W22413 ck: 9557 len: 123 ! Humanised alpha-4 integrin antibody 21.6 VH

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{50}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{57}
1: QVQLVQSGAEVKKPGASVKVSCKASGFNIKDTYIHWVRQAPGQRLEWMGRIDPANGYTKYDPKF

W22428 ck: 9019 len: 142 ! Humanised alpha-4 integrin antibody 21.6 VL

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{69}(I)x(P)(A)xx(Y)(T)(K)xx(P)xxx(G)x{57}
1: MDWTWRVFCLLAVAPGAHSQVQLVQSGAEVKKPGASVKVSCKASGFNIKDTYIHWVRQAPGQRLEW

W26413 ck: 3737 len: 121 ! Swinepox virus HindIII C encoded protein C2

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{83}(A)x(S)(A)xx(F)(T)(H)xx(V)xxx(L)x{22}
1: MIIDVITSLDGGCNKCLNIKHAKSQCKPSSLLISSLENVNPGINPRFFIQNIAANEPEKNIPST

W17971 ck: 6033 len: 47 ! RAC protein kinase C-terminal binding prote

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{23}(L)x(T)(L)xx(L)(C)(R)xx(V)xxx(L)x{8}
1: NSARARASAAPRPGAMLRSCAARLRTLGALCRPPVGRRLPEATRDPS

W34668 ck: 9097 len: 438 ! Arabidopsis thaliana Rec-A protein. DNA enc

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{200}(L)x(T)(A)xx(M)(C)(R)xx(A)xxx(I)x{222}
1: DSQLVLSLKLNPSFTPLSPLFPFTPCSSFSPSLRFSSCYSRRLYSPVTVYAAKKLSHKISSEFDDR
YSKALGVDVENLIVCQPDNGEMALETADRMCRSGAVDLICVDSVSALTPRAEIEGE
IGMQQMGLQARLMSQALRKMSGNASKAGCTLIFLNQIRYKIGVYYGNPEVTSGGIALKFFASVRLEIRSAGKIKSSKGDEDIGLRARV

W24491 ck: 1423 len: 122 ! Novel amylase inhibitor protein subunit isc

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{91}(A)x(S)(I)xx(V)(C)(K)xx(I)xxx(A)x{15}
1: SGPWMCYPGYAFKVPALPGCRPVLKLQCNGSQVPEAVLRDCCQQLADISEWCRCGALYSMLDSM

W24492 ck: 2685 len: 124 ! Novel amylase inhibitor protein 2 subunit i

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{91}(A)x(S)(I)xx(V)(C)(K)xx(I)xxx(A)x{17}
1: SGPWMCYPGYAFKVPALPGCRPVLKLQCNGSQVPEAVLRDCCQQLADISEWCRCGALYSMLDSM

W24493 ck: 3320 len: 124 ! Novel amylase inhibitor protein 3 subunit i

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
x{91}(A)x(S)(I)xx(V)(C)(R)xx(I)xxx(A)x{17}
1: SGPWMCYPGQAFQVPALPACRPLLRLQCNGSQVPEAVLRDCCQQLAHISEWCRCGALYSML

W29523 ck: 2685 len: 124 ! Wheat amylase inhibitor 0.26 Ala subunit

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{91}(A)x(S)(I)xx(V)(C)(K)xx(I)xxx(A)x{17}
1: SGPWMCYPGYAFKVPALPGCRPVLKLQCNGSQVPEAVLRDCCQQLADISEWCRCGALYSML

W29524 ck: 1423 len: 122 ! Wheat amylase inhibitor 0.26 Alb subunit

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{91}(A)x(S)(I)xx(V)(C)(K)xx(I)xxx(A)x{15}
1: SGPWMCYPGYAFKVPALPGCRPVLKLQCNGSQVPEAVLRDCCQQLADISEWCRCGALYSML

W29525 ck: 3320 len: 124 ! Wheat amylase inhibitor 0.19 Al subunit.

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{91}(A)x(S)(I)xx(V)(C)(R)xx(I)xxx(A)x{17}
1: SGPWMCYPGQAFQVPALPACRPLLRLQCNGSQVPEAVLRDCCQQLAHISEWCRCGALYSML

R63906 ck: 7673 len: 248 ! Type I ribosome-inactivating protein luf

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{37}(L)x(S)(A)xx(A)(S)(R)xx(L)xxx(S)x{195}
1: DVRFSLSGSSSTSYSKFIGDLRKALPSNGTVYNLTILLSSASGASRYTLMTLSNYDGKAITVA
QVPSLATISLENSLWSALSKQIQLAQTNNGTFKTPVVIIDDKGQRVEITNVTSKVV
TKNIQLLLNYKQNVA

W45093 ck: 9381 len: 176 ! Residues 217-392 of human type B EBNA2 (

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{155}(P)x(T)(A)xx(V)(C)(R)xx(P)xxx(P)x{5}
1: HQLSLPPHPPPHQSTPHCSSDSTGLPPPPTSYSIPSMTLSPEPLPPPAAPAHPLPGVIYDQQA

W38190 ck: 1840 len: 107 ! Soybean SCARECROW SRPg1 protein. DNA enc

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{70}(I)x(P)(L)xx(F)(S)(H)xx(A)xxx(I)x{21}
1: STENLEDANKMLLEISQLSTPFXTSAQRVAAYFSEAISARLVSSCLGIYATLPHTHQSHKV

W34214 ck: 5054 len: 271 ! Streptomyces ketoreductase (frenolicin g

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{125}(L)x(S)(V)xx(V)(T)(R)xx(L)xxx(G)x{130}
1: MTTAAPHTRPGEAGTTRGPALVTGATRGIGLAVAEALVARGYPVVVCARDAEAVARTVKELAA
TKALGLELARSGITVNAVCPGYVETPMAAGVRRHYADLWDVTEEDVLARFEAKIPL
GRYTRPDEVAALVDYLVTDAAAAVTAQALNVCGGLGNY

W41747 ck: 3783 len: 239 ! Hepatitis C virus antigen. Protein contg

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{84}(L)x(S)(I)xx(A)(C)(K)xx(P)xxx(A)x{139}
1: VCCSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHSMVYSTTSRSASLRQKKVTFDRLQVL
LYDVVSTLPQAVMGPSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTE
NDIRTE

W47606 ck: 9828 len: 194 ! HRSV protein 22K. Production of human re

1 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
x{83}(I)x(S)(I)xx(I)(T)(K)xx(A)xxx(M)x{95}
1: MSRRNPCKFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDKSIDTLSEISGA
ESTVSDTNDHAKNNDTT

```
            W50006  ck: 5567  len: 365   ! Human papillomavirus-18 E2 protein. Treatin

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{61}(V)x(P)(A)xx(I)(S)(K)xx(A)xxx(I)x{288}
            1:      MQTPKETLSERLSCVQDKIIDHYENDSKDIDSQIQYWQLIRWENAIFFAAREHGIQTLNHQVVPAY
SECEKYGNTGTWEVHFGNNVIDCNDSMCSTSDDTVSATQLVKQLQHTPSPYSSTVS
VGTAKTYGQTSAATRPGHCGLAEKQHCGPVNPLLGAATPTGNNKRRKLCSGNTTPIIHLKGDRNSLKCLRYRLRKHSDHYRDISSTWH


            W55590  ck: 428   len: 307   ! H. pylori ORF 09cp10224_429510_c2_46.aa cel

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{187}(A)x(T)(L)xx(A)(S)(R)xx(I)xxx(S)x{104}
            1:      MKKVLLLTLSLSLSFWLHAERNGFYLGLNFAEGSYIQGQGSIGEKASAENALNQAINNAKNSLFPN
YKDMTGRTLDADTLKRASRHIIRKSSGLVIGMELGASTWFASNNLTPFNQVKSRTI
FQLQGKFGVRFSSDEYDIDRYGDENYLGGSSVELGVKVPAFKVNYYSDDYGDKLDYKRVVSVYLNYTYNFKNKH


            W55815  ck: 5054  len: 271   ! Streptomyces roseofulvus frenolicin gene cl

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{125}(L)x(S)(V)xx(V)(T)(R)xx(L)xxx(G)x{130}
            1:      MTTAAPHTRPGEAGTTRGPALVTGATRGIGLAVAEALVARGYPVVVCARDAEAVARTVKELAAGGA
TKALGLELARSGITVNAVCPGYVETPMAAGVRRHYADLWDVTEEDVLARFEAKIPL
GRYTRPDEVAALVDYLVTDAAAAVTAQALNVCGGLGNY


            W56137  ck: 7408  len: 186   ! Open reading frame 2 peptide of bacteriopha

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{74}(L)x(P)(M)xx(F)(S)(K)xx(A)xxx(A)x{96}
            1:      MNPEQLTNTLGSAVSGVAQGLSFLPGIASGVLGYLGAQKQNATAKQIAREQMAFQERMSNTAYQRA
KNLFSRKGR


            W51011  ck: 7216  len: 277   ! Human liver carbonyl reductase. DNA encodin

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{189}(P)x(S)(A)xx(V)(T)(K)xx(V)xxx(S)x{72}
            1:      MSSGIHVALVTGGNKGIGLAIVRDLCRLFSGDVVLTARDVTRGQAAVQQLQAEGLSPRFHQLDID
DTKKGVHQKEGWPSSAYGVTKIGVTVLSRIHARKLSEQRKGDKILLNACCPGWVRT
DMAGPKATKSPEEGAETPVYLALLPPDAEGPHGQFVSEKRVEQW


            W28052  ck: 2273  len: 269   ! Amino acid sequence of phospho-beta-glucosi

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{43}(I)x(P)(V)xx(L)(S)(H)xx(M)xxx(L)x{210}
            1:      MGLKCLRTSIAWTRIFPNGDEDVPNEEGLAFYDRIFDELIAQGIEPVVTLSHFEMPLHLAKHYGGF
AQIANRLRFFFPDVQVRGYYPSYAKKMLARKGYDVGWQEGDDSILQQGTXDYIGFS
YYMSTAVKHDVDTTXENNIVNGGFGFICGESAYRNE


            W37816  ck: 3813  len: 317   ! Human secreted apoptosis-related protein hS

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{170}(P)x(T)(A)xx(V)(T)(K)xx(A)xxx(M)x{131}
            1:      MRAAAAAGGVRTAALALLLGALHWAPARCEEYDYYGWQAEPLHGRSYSKPPQCLDIPADLPLCHTV
TKICAQCEMEHSADGLMEQMCSSDFVVKMRIKEIKIENGDRKLIGAQKKKLLKPG
PLKRKDTKRLVLHMKNGAGCPCPQLDSLAGSFLVMGRKVDGQLLLMAVYRWDKKNKEMKFAVKFMFSYPCSLYYPFFYGAAEPH


            W59024  ck: 9911  len: 71    ! Enterocin-900. Bacterial growth inhibiting

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{27}(L)x(T)(P)xx(L)(S)(K)xx(A)xxx(A)x{28}
            1:        MQNVKELSTKEMKQIIGGENDHRMPNELNTPNNLSKGGAKCGAAIAGGLFGIPKGPLAWAAALA


            W57330  ck: 4443  len: 339   ! Glycerol-3-phosphate dehydrogenase gpsA. Fe

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{68}(L)x(T)(A)xx(A)(S)(R)xx(L)xxx(P)x{255}
            1:      MNQRNASMTVIGAGSYGTALAITLARNGHEVVLWGHDPEHIATLERDRCNAAFLPDVPFPDTL
RVYSNPDFIGVQLGGAVKNVIAIGAGMSDGIGFGANARTALITRGLAEMSRLGAAL
GADPATFMGMAGLGDLVLTCTDNQSRNRRFGMMLGQGMDVQSAQEKIGQVVEGYRNTKEVRELAHRFGVEMPITEEIYQVLYCGK


            W62641  ck: 7060  len: 137   ! Flea serine protease inhibitor variable

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{88}(A)x(T)(A)xx(A)(T)(R)xx(I)xxx(A)x{33}
            1:      EEKLQNVDLQNLTQRMYSVEVILDLPKFKIESEINLNDPLKKLGMSDMFVPGKADFKGLLEGS


            W55052  ck: 4306  len: 118   ! Sunflower antifungal protein MS59 fragme

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{41}(I)x(T)(A)xx(V)(S)(H)xx(A)xxx(C)x{61}
            1:        DPSFPITGEVYTPGNSSFPTVLQNYIRNLRFNETTTPKPFLIITAEHVSHIQAAVVCGKQN


            W48722  ck: 2863  len: 355   ! Human V28 seven transmembrane receptor.

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{80}(V)x(T)(L)xx(W)(T)(H)xx(I)xxx(G)x{259}
            1:      MDQFPESVTENFEYDDLAEACYIGDIVVFGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPK
DYPEVLQEIWPVLRNVETNFLGFLLPLLIMSYCYFRIIQTLFSCKNHKKAKAIKLI
LLVVIVFFLFWTPYNVMIFLETLKLYDFFPSCDMRKDLRLALSVTETVAFSHCCLNPLIYAFAGEKFRRYLYHLYGKCLAVLCGR


            W54728  ck: 3126  len: 24    ! Peptide from Ii p80-103. Increasing upta

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
            1:        LPKPPKPVSKMRMATPLLMQALPM


            W60258  ck: 4443  len: 339   ! Klebsiella pneumoniae glycerol-3-phospha

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{68}(L)x(T)(A)xx(A)(S)(R)xx(L)xxx(P)x{255}
            1:      MNQRNASMTVIGAGSYGTALAITLARNGHEVVLWGHDPEHIATLERDRCNAAFLPDVPFPDTL
RVYSNPDFIGVQLGGAVKNVIAIGAGMSDGIGFGANARTALITRGLAEMSRLGAAL
GADPATFMGMAGLGDLVLTCTDNQSRNRRFGMMLGQGMDVQSAQEKIGQVVEGYRNTKEVRELAHRFGVEMPITEEIYQVLYCGK


            W60267  ck: 843   len: 387   ! Klebsiella pneumoniae DHAT protein. New

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{143}(A)x(T)(A)xx(V)(T)(R)xx(V)xxx(T)x{228}
            1:      MSYRMFDYLVPNVNFFGPNAISVVGERCQLLGGKKALLVTDKGLRAIKDGAVDKTLHYLREAG
INDPLLMIGKPAALTAATGMDALTHAVEAYISKDANPVTDAAAMQAIRLIARNLRQ
AVALGSNLQARENMAYASLLAGMAFNNANLGYVHAMAHQLGGLYDMPHGVANAVLLPHVARYNLIANPEKFADIAELMGENITGL


            W60764  ck: 7422  len: 260   ! Rainbow trout interleukin 1 beta. Rainbo

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{84}(L)x(S)(A)xx(A)(S)(R)xx(A)xxx(S)x{160}
            1:      MDFESNYSLIKNTSESAAWSSKLPQGLDLEVSHHPITMRHIANLIIAMERLKGGEGVTMGTEF
RPTLHLEEVADKDQLKSISQQSDMVRFLFYRRNTGVDISTLESASFRNWFISTDMQ
QDYTKPVDMCQKAAPNRLTTFTIQRHN


            W57889  ck: 5290  len: 271   ! Corn raffinose synthetase. New nucleic a

1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{170}(V)x(P)(V)xx(I)(S)(H)xx(A)xxx(A)x{85}
            1:      QSTHPCAAFHAASRAISGGPIYVSDSVGQHDFALLRRLALPDGTVLRCEGHALPTRDCLFADP
SHERAIKFAPIGLANMLNTAGAVQAFEAKKDASGVTAEVFVKGAGELVAYSSATPR
LCKVNGDEAEFTYKDGVVTVDVPWSGSSSKLCCVQYVY


            W44905  ck: 1135  len: 39    ! "polyproline beta-turn helix" polypeptid
```

```
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{4}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{19}
              1:   AAADPITMFSLTRQVLNVGPRVPIGPNPVLPDQRLPSSA

 W44906  ck: 6449  len: 70    ! "polyproline beta-turn helix" polypeptide D

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{35}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{19}
              1:   AAANPLVLEFTDAGKKANWDGPKSWGLRLYRTGTDPITMFSLTRQVLNVGPRVPIGPNPVLPDQ

 W44907  ck: 3740  len: 127   ! "polyproline beta-turn helix" polypeptide D

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{92}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{19}
              1:   AAAETTGQAYWKPTSSWDLISLKRGNTPWDTGCSKVACGPCYDLSKVSNSFQGATRGGRCNPLV

 W44908  ck: 1205  len: 79    ! "polyproline beta-turn helix" polypeptide P

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{4}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{59}
              1:   AAADPITMFSLTRQVLNVGPRVPIGPNPVLPDQRLPSSPIEIVPAPQPPSPLNTSYPPSTTSTP

 W44909  ck: 4699  len: 110   ! "polyproline beta-turn helix" polypeptide P

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{35}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{59}
              1:   AAANPLVLEFTDAGKKANWDGPKSWGLRLYRTGTDPITMFSLTRQVLNVGPRVPIGPNPVLPDQ

 W44910  ck: 1990  len: 167   ! "polyproline beta-turn helix" polypeptide P

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{92}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{59}
              1:   AAAETTGQAYWKPTSSWDLISLKRGNTPWDTGCSKVACGPCYDLSKVSNSFQGATRGGRCNPLVLE

 W44879  ck: 9655  len: 104   ! "polyproline beta-turn helix" polypeptide P

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{27}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{61}
              1:   EFTDAGKKANWDGPKSWGLRLYRTGTDPITMFSLTRQVLNVGPRVPIGPNPVLPDQRLPSSPIE

 W44882  ck: 8918  len: 61    ! "polyproline beta-turn helix" polypeptide D

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{27}(P)x(T)(M)xx(L)(T)(R)xx(L)xxx(P)x{18}
              1:   EFTDAGKKANWDGPKSWGLRLYRTGTDPITMFSLTRQVLNVGPRVPIGPNPVLPDAAAPHQ

 W62612  ck: 8912  len: 369   ! Human glutamate-binding protein (HGLUBP). N

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{59}(A)x(T)(P)xx(A)(T)(H)xx(P)xxx(G)x{294}
              1:   MKRVFWCLGTTILPPTLDIRGGPSHPCPPMLSLPTLGPLTHSPLSSPPPTVSQGTPMAPAPTPKGA
VSVFTFVAEVKGFVRENVWTYYVSYAVFFISLIVLSCCGDFRRKHPWNLVALSVLT
ASLSYMVGMIASFYNTEAVIMAVGITTAVCFTVVIFSMQTRYDFTSCMGVLLVSMVVLFIFAILCIFIRNRILEIVYASLGALLFTCF

 W30679  ck: 843   len: 387   ! 1,3-Propanediol oxidoreductase. Fermentativ

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{143}(A)x(T)(A)xx(V)(T)(R)xx(V)xxx(T)x{228}
              1:   MSYRMFDYLVPNVNFFGPNAISVVGERCQLLGGKKALLVTDKGLRAIKDGAVDKTLHYLREAGIEV
INDPLLMIGKPAALTAATGMDALTHAVEAYISKDANPVTDAAAMQAIRLIARNLRQ
AVALGSNLQARENMAYASLLAGMAFNNANLGYVHAMAHQLGGLYDMPHGVANAVLLPHVARYNLIANPEKFADIAELMGENITGLSTL

 W69232  ck: 3029  len: 263   ! FcR-II protein sequence. Nucleic acid encod

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{8}(L)x(T)(L)xx(L)(C)(H)xx(I)xxx(V)x{239}
              1:   MALVLILQLLTLWPLCHTDITPSVPPASYHPKPWLGAQPATVVTPGVNVTLRCRAPQPAWRFG
PWADFTLLGARAPGTYSCYYHTPSAPYVLSQRSEVLVISWEDSGSSDYTRGNLVRL
GLAGLVLISLGALVTFDWRSQNRAPAGIRP

 W62617  ck: 5610  len: 212   ! Rattus norvegicus SOCS1 protein. Suppres

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{171}(V)x(P)(L)xx(L)(C)(R)xx(I)xxx(V)x{25}
              1:   MVARNQVEADNAISPASEPRRRPEPSSSSSSSSPAAPARPRPCPVVPAPAPGDTHFRTFRSHS
LCRQRIVAAVGRENLARIPLNPVLRDYLSSFPFQI

 W62622  ck: 4000  len: 130   ! Mus musculus SOCS8 protein. Suppressor o

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{84}(P)x(S)(L)xx(L)(C)(R)xx(L)xxx(L)x{30}
              1:   MSAILKVGHHCWLPVTSAVNPQRMLRPPPTAVFNCAACCCLWGQMLMNTYRVVQLPEEAKGLV

 W62623  ck: 6307  len: 207   ! Homo sapiens SOCS11 protein. Suppressor

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{120}(V)x(S)(L)xx(L)(C)(R)xx(I)xxx(V)x{71}
              1:   LEKCGWYWGPMNWEDAEMKLKGKPDGSFLVRDSSDPRYILSLSFRSQGITHHTRMEHYRGTFS
PNRSKRWNPPRSEGLPAGHHQGHLVAKLQL

 W62624  ck: 8778  len: 134   ! Mus musculus SOCS13 protein. Suppressor

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{94}(P)x(P)(L)xx(L)(C)(R)xx(V)xxx(L)x{24}
              1:   GGWDLGRNRLYHDGKNQPSKTYPAFLEPDETFIVPDSFFVALDMXDGTLSFIVDGQYMGVAFR

 W62613  ck: 4527  len: 212   ! Mus musculus SOCS1 protein. Suppressor o

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{171}(V)x(P)(L)xx(L)(C)(R)xx(I)xxx(V)x{25}
              1:   MVARNQVAADNAISPAAEPRRRSEPSSSSSSSSPAAPVRPRPCPAVPAPAPGDTHFRTFRSHS
LCRQRIVAAVGRENLARIPLNPVLRDYLSSFPFQI

 W62614  ck: 1060  len: 198   ! Mus musculus SOCS2 protein. Suppressor o

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{159}(A)x(T)(L)xx(F)(C)(R)xx(I)xxx(T)x{23}
              1:   MTLRCLEPSGNGADRTRSQWGTAGLPEEQSPEAARLAKALRELSQTGWYWGSMTVNEAKEKLK
IWGLPLPTRLKDYLEEYKFQV

 W62615  ck: 9828  len: 225   ! Mus musculus SOCS3 protein. Suppressor o

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{185}(V)x(T)(L)xx(L)(C)(R)xx(V)xxx(L)x{24}
              1:   MVTHSKFPAAGMSRPLDTSLRLKTFSSKSEYQLVVNAVRKLQESGFYWSAVTGGEANLLLSAE
LSRPLSSNVATLQHLCRKTVNGHLDSYEKVTQLPGPIREFLDQYDAPL

 W62616  ck: 758   len: 211   ! Homo sapiens SOCS1 protein. Suppressor o

1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{170}(V)x(P)(L)xx(L)(C)(R)xx(I)xxx(V)x{25}
              1:   MVAHNQVAADNAVSTAAEPRRRPEPSSSSSSSPAAPARPRPCPAVPAPAPGDTHFRTFRSHAD
CRQRIVATVGRENLARIPLNPVLRDYLSSFPFQI

 W61619  ck: 1162  len: 250   ! Clone HTPEF86 of TM4SF superfamily. New

 1               <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
```

```
                x{23}(P)x(T)(P)xx(M)(S)(H)xx(L)xxx(S)x{211}
          1:    MNSMTSAVPVANSVLVVAPHNGYPVTPGIMSHVPLYPNSQPQVHLVPGNPPSLVSNVNGQPVQKAL
VNPGMAISGVLLVFCLLEFGIACASSHFGCQLVCCQSSNVSVIYPNIYAANPVITP
EPVTSPPSYSSEIQANK

          W65460  ck: 1814  len: 375   ! Human growth differentiation factor-8. Tran

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSKLRL
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFLQKY

          W64373  ck: 3575  len: 120   ! Mycobacterium tuberculosis antigen RDIF5. N

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{95}(A)x(S)(A)xx(A)(C)(R)xx(P)xxx(P)x{9}
          1:        TPEKSFVDDLDIDSLSMVEIAVQTEDKYGVKIPDEDLAGLRTVGDVVAYIQKLEEENPEAAQAL

          W38580  ck: 5404  len: 174   ! Streptococcus pneumoniae protein of unknown

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{72}(V)x(T)(V)xx(L)(S)(K)xx(V)xxx(I)x{86}
          1:      MGFNQPLSWLILHAKGLTESFIKASAQTGADRITYISCNVRNHGAVDIKLYQELGYELKKVQPVDL

          W70962  ck: 4556  len: 212   ! A STAT function regulatory protein designat

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{171}(V)x(P)(L)xx(L)(C)(R)xx(I)xxx(V)x{25}
          1:      MVARNQVAADNAISPAAEPRRRSEPSSSSSSSSPAAPVRPRPCPGVPAPAPGDTHFRTFRSHSDYR
LCRQRIVAAVGRENLARIPLNPVLRDYLSSFPFQI

          W77666  ck: 5825  len: 100   ! Staphylococcus aureus protein of unknown fu

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{76}(A)x(T)(P)xx(A)(S)(K)xx(I)xxx(I)x{8}
          1:        VMPXVFWXPFPXIFIGTALPLAGTVATGAIHFTANEVIPIGXMLXNNGLIAINLAYQNLDRAFV

          W38716  ck: 3773  len: 161   ! S. pneumoniae aspartyl tRNA synthetase. Nov

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{53}(P)x(T)(A)xx(Y)(S)(R)xx(I)xxx(T)x{92}
          1:      MKYDDAMALYGSDKPDTRFDMLLQDLTEVVKGVDFKVFSEALAVKAIVVNGLEPLTADNYSRKDID

          W79302  ck: 8790  len: 67    ! A Staphylococcus aureus protein. New nuclei

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{50}(L)x(S)(A)xx(I)(T)(K)xx(I)xxx(M)x
          1:        SHNLNMRGAFLHVIGDLLGSVGAITAAILIWAFGWTIADPIASILVSVIILKSAWGITKSSINI

          W69889  ck: 1496  len: 376   ! Rat growth differentiation factor-8. Transg

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{66}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MIQKPQMYVYIYLFVLIAAGPVDLNEDSEREANVEKEGLCNACAWRQNTRYSRIEAIKIQILSKLR
IKPMKDGTRYTGIRSLKLDMSPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDE
NGHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFLQK

          W69891  ck: 1805  len: 375   ! Pig growth differentiation factor-8. Transg

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                    x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MQKLQIYVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACMWRQNTKSSRLEAIKIQILSKLRL
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFL

          W69892  ck: 1229  len: 375   ! Ovine growth differentiation factor-8. T

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MQKLQIFVYIYLFMLLVAGPVDLNENSEQKENVEKKGLCNACLWRQNNKSSRLEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGIGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPEPGEEGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFLFL

          W30689  ck: 2293  len: 376   ! Murine growth differentiation factor-8.

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{66}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MMQKLQMYVYIYLFMLIAAGPVDLNEGSEREENVEKEGLCNACAWRQNTRYSRIEAIKIQILS
IKPMKDGTRYTGIRSLKLDMSPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDE
NGHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVF

          W69885  ck: 1814  len: 375   ! Human growth differentiation factor-8. T

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFL

          W69886  ck: 1463  len: 375   ! Baboon growth differentiation factor-8.

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEALGWDWIIAPKRYKANYCSGECEFVFL

          W69887  ck: 9305  len: 375   ! Bovine growth differentiation factor-8.

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
          1:      MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFL

          W80799  ck: 2018  len: 244   ! Rhodococcus nitrile hydratase gene fragm

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{179}(A)x(P)(M)xx(F)(S)(R)xx(L)xxx(V)x{49}
          1:      MAGADVHAQGGTNRRARILVVDDEKHVRTMVTWQLESENFDVVAAADGDAALRQVTESAPDLM
MAASPMQVFSRRRLLLEVWRSSPDWQQDATVTEHVHRIRRKIEEDPTKPTILQTVR
GAGYRFDGERA

          W80624  ck: 4504  len: 189   ! S. pneumoniae GTP binding protein. Strep

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{67}(I)x(P)(I)xx(L)(T)(K)xx(L)xxx(L)x{106}
          1:      MSASEGRDPYEDYLAINKELESYNLRLMERPQIIVTNKMDMPESQENLEEFKKKLAENYDEFE
VRIGKFEFEFVD

          W68494  ck: 9982  len: 368   ! E2 papillomavirus protein encoded by pla

1                 <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                    x{62}(V)x(P)(A)xx(I)(S)(K)xx(A)xxx(I)x{290}
          1:      TMETPKETLSERLSCVQDKIIDHYENDSKDIDSQIQYWQLIRWENAIFFAAREHGIQTLNHQV
KSECEKYGNTGTWEVHFGNNVIDCNDSMCSTSDDTVSATQLVKQLQHTPSPYSSTV
SVGTAKTYGQTSAATRPGHCGLAEKQHCGPVNPLLGAATPTGNNKRRKLCSGNTTPIIHLKGDRNSLKCLRYRLRKHSDHYRDIS
```

```
W82504  ck: 447   len: 225   ! Human EPRG1 protein #1. New EPO primary res
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{185}(V)x(T)(L)xx(L)(C)(R)xx(V)xxx(L)x{24}
             1:  MVTHSKFPAAGMSRPLDTSLRLKTFSSKSEYQLVVNAVRKLQESGFYWSAVTGGEANLLLSAEPAG
LSRPLSSNVATLQHLCRKTVNGHLDSYEKVTQLPGPIREFLDQYDAPL

W82505  ck: 9756  len: 157   ! Human EPRG1 protein #2. New EPO primary res
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{117}(V)x(T)(L)xx(L)(C)(R)xx(V)xxx(L)x{24}
             1:  EFANSSDQRHFFQLSVKTQSGTKNLRIQCEGGSFCLQSDPRKTQPVPRVDCVLKLVHRSGAPAGAF

W81740  ck: 3575  len: 120   ! M. tuberculosis immunogenic polypeptide RDI
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{95}(A)x(S)(A)xx(A)(C)(R)xx(P)xxx(P)x{9}
             1:  TPEKSFVDDLDIDSLSMVEIAVQTEDKYGVKIPDEDLAGLRTVGDVVAYIQKLEEENPEAAQAI

W72902  ck: 5027  len: 226   ! Mycobacterium tuberculosis antigen CFP19A.
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{112}(M)x(S)(A)xx(A)(T)(R)xx(L)xxx(S)x{98}
             1:  MIPRPQPHSGRWRAGAARRLTSLVAAAFAAATLLLTPALAPPASAGCPDAEVVFARGTGEPPGLGR
ALTPQFGSKTINLCNNGDPICSDGNRWRAHLGYVPGMTNQAARFVASRI

W80495  ck: 1679  len: 207   ! Human vascular endothelial growth factor (V
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{158}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  MSPLLRRLLLAALLQLAPAQAPVSQPDAPGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGTVAK
AAPSTTSALTPGPAAAAADAAASSVAKGGA

W85306  ck: 3126  len: 24    ! Helper T-cell class II peptide derived from
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
             1:    LPKPPKPVSKMRMATPLLMQALPM

W86202  ck: 9273  len: 206   ! Human VEGF-related factor (VRF)-2 sequence.
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{157}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  MSPLLRRLLLAALLQLAPAQAPVSQPDAPGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGTVAK
APSTTSALTPGPAAAAADAAASSVAKGGA

W68541  ck: 8848  len: 420   ! Amino acid sequence of the SIAX DP2-64 (Oct
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{193}(M)x(S)(L)xx(L)(S)(H)xx(A)xxx(M)x{211}
             1:  MMSMNSKQPHFAMHPTLPEHKYPSLHSSSEAIRRACLPTPPLQSNLFASLDETLLARAEALAAVDI
GPGGGLLGGSAHPHPHMHSLGHLSHPAAAAAMNMPSGLPHPGLVAAAAHHGAAAAA
AAAAAGQVAAASAAAAVVGAAGLASICDSDTDPRELEAFAERFKQRRIKLGVTQADVGSALANLKIPGVGSLSQSTICRFESLTLSHN

W86235  ck: 8227  len: 185   ! Human VRF (VEGF-related factor)-2 full leng
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                   x{136}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  PVSQPDAPGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGGCCPDDGL
SSVAKGGA

W86214  ck: 5797  len: 178   ! Human VRF-2 truncated fragment 1. New trunc
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{129}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  PGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGGCCPDDGLECVP
A

W86215  ck: 4363  len: 173   ! Human VRF-2 truncated fragment 2. New tr
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{124}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  KVVSWIDVYTRATCQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGGCCPDDGLECVPTGQHQ

W86216  ck: 2203  len: 168   ! Human VRF-2 truncated fragment 3. New tr
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{119}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  IDVYTRATCQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGGCCPDDGLECVPTGQHQVRMQI

W86217  ck: 3123  len: 163   ! Human VRF-2 truncated fragment 4. New tr
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{114}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
             1:  RATCQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGGCCPDDGLECVPTGQHQVRMQILMIRY

W85119  ck: 5663  len: 446   ! A delta-5 desaturase enzyme. New nucleic
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{166}(L)x(P)(L)xx(A)(S)(H)xx(V)xxx(P)x{264}
             1:  MGTDQGKTFTWEELAAHNTKDDLLLAIRGRVYDVTKFLSRHPGGVDTLLLGAGRDVTPVFEMY
VTHNPTVWKILGATHDFFNGASYLVWMYQHMLGHHPYTNIAGADPDVSTSEPDVRR
IKPNQKWFVNHINQHMFVPFLYGLLAFKVRIQDINILYFVKTNDAIRVNPISTWHTVMFWGGKAFFVWYRLIVPLQYLPLGKVLL

W73507  ck: 3945  len: 317   ! Human ATG-1709 protein. New human secret
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{170}(P)x(T)(A)xx(V)(T)(K)xx(A)xxx(M)x{131}
             1:  MRAAAAAGGVRTAALALLLGALHWAPARCEEYHYYGWQAEPLHGRSYSKPPQCLDIPADLPLC
TKICAQCEMEHSADGLMEQMCSSDFVVKMRIKEIKIENGDRKLIGAQKKKKLLKPG
PLKRKDTKRLVLHMKNGAGCPCPQLDSLAGSFLVMGRKVDGQLLLMAVYRWDKKNKEMKFAVKFMFSYPCSLYYPFFYGAAEPH

W90011  ck: 9038  len: 343   ! Expressed antigen for cluster 35a. New H
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{42}(I)x(S)(I)xx(I)(S)(K)xx(V)xxx(G)x{285}
             1:  MATKLTPKQKAQLDELSMSEKIAILLIQVGEDTTGEILRHLDIDSITEISKQIVQLNGTDKQI
EISPQVVKRVSTVLENKLESLTSYKIEVGGLRAVAEIFNRLGQKSAKTTLARIESV
DNKLAGAIKEMMFTFEDIVKLDNFAIREILKVADKKDLSLALKTSTKDLTDKFLNNMSSRAAEQFVEEMQYLGAVKIKDVDVAQR

W89910  ck: 8829  len: 343   ! Antigen 1 from cluster 35a. New Helicoba
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{42}(I)x(S)(I)xx(I)(S)(K)xx(V)xxx(G)x{285}
             1:  MATKLTPKQKAQLNELSMSEKIAILLIQVGEDTTGEILRHLDIDSITEISKQIVQLNGTDKQI
EISPQVLKRVSTVLENKLESLTSYKIEVGGLRAVAEIFNRLGQKSAKTTLARIESV
DNKLAGAIKEMMFTFEDIVKLDNFAIREILKVADKKDLSLALKTSTKDLTDKFLNNMSSRAAEQFVEEMQYLGAVKIKDVDVAQR

W85590  ck: 2306  len: 354   ! Polypeptide with 98% homology with human
1                <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                   x{115}(A)x(S)(A)xx(I)(S)(H)xx(A)xxx(T)x{223}
             1:  MRARPQVCEALLFALALQTGVCYGIKWLALSKTPSALALNQTQHCKQLEGLVSAQVQLCRSNL
VKKTGSQANKLMRLHNSEVGRQALRASLEMKCKCHGVSGSCSIRTCWKGLQELQDV
AADLKTRYLSATKVVHRPMGTRKHLVPKDLDIRPVKDSELVYLQSSPDFCMKNEKVGSHGTQDRQCNKTSNGSDSCDLMCCGRGY

W85591  ck: 8607  len: 351   ! Polypeptide with 95% homology with human
```

```
1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{112}(A)x(S)(A)xx(I)(S)(H)xx(A)xxx(T)x{223}
                 1:  RPQVCEAMLFALALQTGVCYGIKWLALSKTPSALALNQTQHCKQLEGLVSAQVQLCRSNLELMHTV
TGSQANKLMGLHNSEVGRQALRASLEMKCKCHGVSGHCYIGTSWKGLQELQDVAAD
LKTRYLSATKVVHRPMGTRKHLVPKDLDIRPVKDSELVYLQSSPDFCMKNEKVGSHGTQDRQCNKTSNGSDSCDLMCCGRGYNPYTDR

W89100  ck: 1342  len: 36    ! Polypeptide fragment encoded by gene 206. N

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{5}(P)x(T)(L)xx(I)(S)(R)xx(L)xxx(P)x{15}
                 1:  FSVHRPETLFNISRFLLHSLPKDTPSGISKVKILFT

W88568  ck: 9823  len: 221   ! Secreted protein encoded by gene 35 clone H

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{160}(M)x(T)(M)xx(M)(S)(H)xx(P)xxx(G)x{45}
                 1:  MFHGIPATPGIGAPGNKPELYEEVKLYKNARE REKYDNMAELFAVVKTMQALEKAYIKDCVSPSEY
QTVSQWLQTLSGMSASDELDDSQVRQMLFDLESAYNAFNRFLHA

W83391  ck: 9212  len: 161   ! Caenorhabditis elegans synMuv protein LIN-5

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{88}(A)x(T)(V)xx(M)(S)(H)xx(A)xxx(A)x{57}
                 1:  MSRPLGFIGYEFGDDEMFVQQMIEKKSNAEQAKMLEQQKKMLECTETMPEESEPVPMKCLDFEEAF

W95506  ck: 5663  len: 446   ! Mortierella alpina delta 5 desaturase. Nucl

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{166}(L)x(P)(L)xx(A)(S)(H)xx(V)xxx(P)x{264}
                 1:  MGTDQGKTFTWEELAAHNTKDDLLLAIRGRVYDVTKFLSRHPGGVDTLLLGAGRDVTPVFEMYHAF
VTHNPTVWKILGATHDFFNGASYLVWMYQHMLGHHPYTNIAGADPDVSTSEPDVRR
IKPNQKWFVNHINQHMFVPFLYGLLAFKVRIQDINILYFVKTNDAIRVNPISTWHTVMFWGGKAFFVWYRLIVPLQYLPLGKVLLLFT

W98862  ck: 4789  len: 304   ! H. pylori GHPO 1670 protein. New isolated H

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x(L)x(S)(A)xx(Y)(C)(K)xx(A)xxx(L)x{287}
                 1:  MLESALKYCKEKAIDLLVGFVPKTYSMAQECNILGLYDDAFIITKQENLVGIISLQGLSYSNLMQK
NTSSKKDKKNKFKKKITFSTKSKRALLIQTIERVKNALKEFKPTLLNSKEVLNFYA
EYINGKYIAFNPKLKRLSDTILHLMCILRKITLSLNFKIKTPFVRVWGLRLMRAKKFLRSLYLLFYTPKLN

W98755  ck: 3203  len: 102   ! H. pylori GHPO 1075 protein. New isolated H

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{66}(I)x(T)(I)xx(F)(S)(K)xx(P)xxx(C)x{20}
                 1:  MIFSFKRTSFCVSVPVLSIKIFFTLARVSKNNASLNTIKGFLNTPITNAMGVARANAQGQLMTS

W98226  ck: 2117  len: 115   ! H. pylori GHPO 1099 protein. New isolated H

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{13}(V)x(S)(L)xx(A)(C)(K)xx(P)xxx(S)x{86}
                 1:  MKSKITHFIAISFVLSLFSACKDEPKKSSQSHQNNTKITKNNPINQANNDIRKIEHEEEDEKAT

W88330  ck: 7734  len: 336   ! Mannosyl transferase involved in O antigen

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                       x{115}(I)x(T)(V)xx(F)(S)(K)xx(I)xxx(S)x{205}
                 1:  MNNKKVLMDISWSNKGGIGRFTDEISKLLCDISKEELYRKCASPLAPLGLAVNIFLRKKTDVVFLF
VISAFAKADIDPSIKLVFTGNPCNDLEKLIIQHGLSERVKFFGFVSEKDLPSLYKG
SLGLVFPSLYEGFGLPVVEGMACGIPVLTSLTSSLPEVAGDAAILVDPLSEDAITKGISRLINDSELRKHLIQKGLLRAKRFNWQNVV

Y01440  ck: 9731  len: 64    ! Secreted protein encoded by gene 58 clone H

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x(A)x(S)(V)xx(A)(T)(R)xx(A)xxx(T)x{47}
                 1:  MAASVGRATRSAAAHLTQLPPAPRAQRTSPAQPDEGKRRDADPWRTGPTVNKTGSIPGRLR

W99185  ck: 9793  len: 314   ! Rhodococcus corallina ohp operon protein

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{141}(A)x(P)(L)xx(V)(S)(R)xx(L)xxx(A)x{157}
                 1:  MPVALCAMSHSPLMGRNDPEQEVIDAVDAAFDHARRFVADFAPDLIVIFAPDHYNGVFYDLLP
SHDPPVPQFATAPEEVRERLIDGRNPSAAERDAREQRVITAGRDFAAGTAAIQPLN
PEWDRHLLDVLASGDLEQIDAWTNDWFVEQAGHSSHEVRTWIAAYAAMSAAGKYRVTSTFYREIHEWIAGFGITTAVAVDE

W97722  ck: 1929  len: 302   ! Staphylococcus aureus mutant P15B9 virul

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{24}(L)x(T)(A)xx(A)(S)(K)xx(I)xxx(M)x{262}
                 1:  MNQKGRGLEILINEKQDGQWLFSVLKTALKASKPVIQDWMSHQQIKVNHESVLNNMIVKKGDR
LRTKKGTINPPIGRDRSHPTRRRVSPGGQTAVTHFKVMASNAKERLSLVELELETG
RTHQIRVHLASLGHPLTGDSLYGGGSKLLNRQALHANKVQAVHPITDELIVAEAPFPADMKNLCRTYFS

Y11009  ck: 2840  len: 142   ! H. pylori ORF hp6p22217_23564012_f1_5 se

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{42}(L)x(S)(V)xx(V)(S)(R)xx(L)xxx(I)x{84}
                 1:  MKTNFYKIKLLFAWCLIIGMFNAPLNADQNTDIKDISPEDMALNSVGLVSRDQLKIEIPKETL

Y11062  ck: 4555  len: 116   ! H. pylori ORF 06cp30603_10744075_c3_136

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{14}(V)x(S)(V)xx(A)(C)(K)xx(P)xxx(S)x{86}
                 1:  LMKSKITHFIVISFVLSVLSACKDEPKKSSQSHQNNTKTTQNNQINQPNKDIKKIEHEEEED

W97887  ck: 1814  len: 375   ! Human myostatin. Increasing muscle mass

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
                 1:  MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFL

W97884  ck: 8371  len: 375   ! Bovine myostatin. Increasing muscle mass

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
                 1:  MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGICQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFL

W97885  ck: 1462  len: 286   ! Bovine myostatin (mutant form). Increasi

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{65}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
                 1:  MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSK
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDEN
GHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRRDFGLDCDRISMLSLPSNCGF

W97886  ck: 2420  len: 376   ! Murine myostatin. Increasing muscle mass

1                    <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                       x{66}(L)x(T)(A)xx(I)(S)(K)xx(I)xxx(L)x{294}
                 1:  MMQKLQMYVYIYLFMLIAAGPVDLNEGSEREENVEKEGLCNACAWRQNTRYSRIEAIKIQILS
IKPMKDGTRYTGIRSLKLDMSPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDE
NGHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRRDFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVF
```

```
            W97840  ck: 7697  len: 207   ! Winged bean chymotrypsin inhibitor WCI-3. A
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{69}(V)x(S)(P)xx(V)(S)(K)xx(P)xxx(S)x{122}
                1:    MKSTTFLALFLLSAIISHLPSSTADDDLVDAEGNLVENGGTYYLLPHIWAHGGGIETAKTGNEPCP
RNGNRRLVVTEENPLELVLLKAKSETASSH

            Y03794  ck: 7277  len: 144   ! S. aureus polypeptide. New essential polype
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{17}(L)x(T)(L)xx(I)(T)(K)xx(V)xxx(G)x{111}
                1:    MDRTKQSLNVFVGMNRALDTLEQITKEDVKRYGLNITEFAVLELLYNKGPQPIQRIRDRVLIASSS

            Y03782  ck: 881   len: 135   ! S. aureus polypeptide. New essential polype
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{92}(I)x(S)(I)xx(L)(T)(R)xx(L)xxx(S)x{27}
                1:    IRHERAFIKRGLLQFADGNGFQCGIWYEGTLVGVIGLHEINHMHRKTSLGYYLYKQFEGHGIMTQA

            Y02934  ck: 2639  len: 53    ! Fragment of human secreted protein encoded
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{3}(A)x(S)(P)xx(L)(S)(R)xx(M)xxx(C)x{34}
                1:    TLVAGSPCSLSRWIMAGFCHGELVQSDMESQEWERGQVVLSHTSLPWCYVSPR

            Y11970  ck: 1817  len: 113   ! Human 5' EST secreted protein SEQ ID No: 57
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{9}(V)x(P)(L)xx(L)(S)(R)xx(V)xxx(A)x{88}
                1:    MANLFIRKMVNPLLYLSRHTVKPRALSTXLFGSIRGAAPVAVEPGAAVRSLLSPGLLPHLLPAL

            W96262  ck: 3324  len: 423   ! Brn-3a polypeptide. New polypeptide compris
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{196}(M)x(S)(L)xx(L)(S)(H)xx(A)xxx(M)x{211}
                1:    MMSMNSKQPHFAMHPTLPEHKYPSLHSSSEAIRRACLPTPPLQSNLFASLDETLLARAEALAAVDI
GGGGPGGGLLGGSAHPHPHMHSLGHLSHPAAAAAMNMPSGLPHPGLVAAAAHHGAA
AAAAAAAAGQVAAASAAAAVVGAAGLASICDSDTDPRELEAFAERFKQRRIKLGVTQADVGSALANLKIPGVGSLSQSTICRFESLTL

            Y05204  ck: 3125  len: 24    ! Human CLIP immunomodulatory peptide. Use of
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x(P)x(P)(P)xx(V)(S)(K)xx(M)xxx(L)x{7}
                1:    KPKPPKPVSKMRMATPLLMQALPM

            Y11993  ck: 4119  len: 102   ! Human 5' EST secreted protein SEQ ID No: 59
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{80}(V)x(T)(L)xx(W)(T)(H)xx(I)xxx(G)x{6}
                1:    MDQFPESVTENFEYDDLAEACYIGDIVVFGTVFLSIFYSVIFAIGLVGNLLVVFALTNSKKPKS

            Y05545  ck: 5919  len: 243   ! Wheat Type III glutathione transferase subu
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{91}(I)x(P)(A)xx(Y)(S)(R)xx(A)xxx(A)x{136}
                1:    MAGGEELKLLGWWAPGVSPYVLRAQMALAVKGLSYDYLPEDRWSTSDLLIASNPVYKKVPVLIHNG
PWFEAAGRLAGLGPIIDPARTPKLAAWAERFSVAEPIKALLPGVDKLEEYITTALY
PKWNIAVTGN

            Y07041  ck: 1679  len: 207   ! Breast cancer associated antigen precursor
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W)(C
                      x{158}(A)x(S)(P)xx(I)(T)(H)xx(P)xxx(P)x{33}
                1:    MSPLLRRLLLAALLQLAPAQAPVSQPDAPGHQRKVVSWIDVYTRATCQPREVVVPLTVELMGTVAK
AAPSTTSALTPGPAAAAADAAASSVAKGGA

            Y07007  ck: 4692  len: 177   ! Breast cancer associated antigen precurs
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{47}(V)x(S)(P)xx(I)(S)(K)xx(A)xxx(G)x{114}
                1:    LCCMHYCCKSCWNEYLTTRIEQNLVLNCTCPIADCPAQPTGAFIRAIVSSPEVISKYKALLRG

            W93953  ck: 7759  len: 352   ! Human regulatory molecule HRM-9 protein.
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{17}(P)x(S)(P)xx(L)(T)(K)xx(L)xxx(A)x{319}
                1:    MHVVAPASLRLGTGTNLPPSPTCLTKLALPPAAEPSLLAMSQSRHRAEAPPLEREDSGTFSLG
GWRLPEYTLSQEGGPAHKREYTTICRLESFMETGKGASKKQAKRNAAEKFLAKFSN
ISPENHISLTNVVGHSLGCTWHSLRNSPGEKINLLKRSLLSIPNTDYIQLLSEIAKEQGFNITYLDIDELSANGQYQCLAELSTS

            Y04959  ck: 8454  len: 273   ! Mycobacterium species protein sequence 4
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{185}(P)x(T)(V)xx(W)(S)(H)xx(A)xxx(L)x{72}
                1:    MANKRGNAGQPLPLSDRDDDHMQGHWLLARLGKRVLRPGGVELTRTLLARAEVTDADVLELAP
RALKVNARPLTVAEWSHLLAGHGLVVEHVVTASMALLQPRRVIADEGLLGALRFAG
NLLIHRAARRRVLLMRHTFRRHRERLTAVAIVAHKPHVDS

            Y04960  ck: 8320  len: 280   ! Mycobacterium species protein sequence 4
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{192}(P)x(T)(V)xx(W)(S)(H)xx(A)xxx(L)x{72}
                1:    KPAEVGSMANKRGNAGQPLPLSDRDDDHMQGHWLLARLGKRVLRPGGVELTRTLLARAEVTDA
DLRQSLARALKVNARPLTVAEWSHLLAGHGLVVEHVVTASMALLQPRRVIADEGLL
GALRFAGNLLIHRAARRRVLLMRHTFRRHRERLTAVAIVAHKPHVDS

            Y04901  ck: 6286  len: 72    ! Mycobacterium species protein sequence 3
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{50}(P)x(S)(I)xx(A)(T)(R)xx(A)xxx(S)x{6}
                1:    ANTREGSATAALPAPWLLNNTSQPRTLPVCGRINDPNSTNTRIATTKALAPGSISRATRRS

            Y04743  ck: 5429  len: 110   ! Mycobacterium species protein sequence 1
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{14}(L)x(T)(L)xx(A)(S)(R)xx(V)xxx(A)x{80}
                1:    MAGDQELELRFDVPLYTLAEASRYLVVPRATLATWADGYERRPANAPAVQGQPIAFDAYSV

            Y07239  ck: 2144  len: 141   ! Fragment of human placental STAT3 protei
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{70}(I)x(S)(V)xx(Y)(T)(K)xx(L)xxx(S)x{55}
                1:    WLDNIIDLVKKYILALWNEGYIMGFISKERERAILSTKPPGTFLLRFSESSKEGGVTFTWVEK

            Y07240  ck: 2614  len: 141   ! Fragment of human hepatic STAT3 protein
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{70}(I)x(S)(V)xx(Y)(T)(K)xx(L)xxx(S)x{55}
                1:    WLDNIIDLVKKYILALWNEGYIMGFISKERERAILSTKPPGTFLLRFSESSKEGGVTFTWVEK

            Y07241  ck: 3064  len: 141   ! Fragment of mouse hepatic STAT3 protein
1                   <(X){1,200}(L,I,V,M,A,P)X(P,T,S)(L,I,V,M,A,P)XX(L,I,V,M,A,F,Y,W
                      x{70}(I)x(S)(V)xx(Y)(T)(K)xx(L)xxx(S)x{55}
                1:    WLDNIIDLVKKYILALWNEGYIMGFISKERERAILSTKPPGTFLLRFSESSKEGGVTFTWVEK
```

Databases searched:
GeneSeq, Release 36.3, Released on 8Jul1999, Formatted on 9Jul1999

Total finds: 245
Total length: 23,686,106
Total sequences: 188,963
CPU time: 03:36.13

!!AA_SEQUENCE 1.0
ID P90009 standard; protein; 102 AA.
AC P90009;
DT 1-NOV-1989 (first entry)
DE Mouse gamma-2b chain
KW Immunoglobulin E competitor; low affinity receptor sites;
KW Bennich enumeration; reduced hypersensitivity.
OS Murine
PN WO8904834-A.
PD 01-JUN-1989.
PF 18-NOV-1988; G01018.
PR 19-NOV-1987; GB-027045.
PA (RECO) Research Corp Ltd.
PI Gould HJ, Helm BA, Marsh PJHB;
DR WPI; 89-178361/24.
DR N-PSDB; N90237.
PT New immunoglobulin E competitor
PT - binding exclusively with low affinity receptors (FcER2)
PS Disclosure; Table 4; 33pp; English.
CC Amino acid residues (see corresp. DNA N90237) representing the
CC mouse gamma-2b chain. Used to extend P90006 (as Region_2) to
CC create an IgE competitor binding to type 2 low affinity
CC Fc receptors.
SQ Sequence 102 AA;

P90009 Length: 102 February 11, 2000 15:48 Type: P Check: 7203 ..

1 PQVYILPPPA EQLSRKDVSL TCLVVGFNPG DISVEWTSNG HTEENYKDTA

51 PVLDSDGSYF IYSKLDIKTS KWEKTDSFSC NVRHEGLKNY YLKKTISRSP

101 GK

!!AA_SEQUENCE 1.0
ID P91884 standard; peptide; 48 AA.
AC P91884;
DT 26-APR-1990 (first entry)
DE Antigenic Epstein-Barr virus peptide
KW Eppstein-Barr virus; EBV; EBV antigen; EBV vaccine; EBV peptide.
OS Epstein-Barr virus
PN EP-316170-A.
PD 17-MAY-1989.
PF 10-NOV-1988; 310591.
PR 11-NOV-1987; IE-003041.
PA (GGEN-) G Gene Galway Ltd.
PI Gannon BFX, Walls DFG;
DR WPI; 89-146691/20.
DR N-PSDB; N91751.
PT Identifying antigens from a pathogenic organism expressed in vivo
PT - by generating a gene bank using nucleic acid
PT fragments and screening with infected serum
PS Claim 8(3); 21pp; English.
CC It is coded for by a DNA sequence isolated by using sera from patients
CC infected with EBV to screen colonies of micro-organisms transformed with
CC fragments of EBV DNA.
SQ Sequence 48 AA;

P91884 Length: 48 February 11, 2000 15:48 Type: P Check: 718 ..

1 RGWFCPSLCP SEEPGTSGTP EPLGPASRRP PGLRSPLSPV KPKECLRG

!!AA_SEQUENCE 1.0
ID P94157 standard; protein; 246 AA.
AC P94157;
DT 12-JUN-1990 (first entry)
DE Acetyl-CoA reductase.
KW Polyester biopolymers; acetyl-CoA reductase; polyhydroxybutyrate; ss.
OS Alcaligenes eutrophus H16.
PN WO8900202-A.
PD 12-JAN-1989.
PF 27-JUN-1988; U02227.
PR 29-JUN-1987; US-067695.
PA (MASI) Massachusetts Institute of Technology.
PI Peoples OP, Sinskey AJ;
DR WPI; 89-039655/05.
DR N-PSDB; N91209.
PT Constructing new polyester biopolymers - using genes encoding
PT beta-ketothiolase(s), acetoacetyl-CoA reductase(s) and
PT polyhydroxybutyrate synthetase(s).
PS Disclosure; p; English.
CC The enzyme is used to study the polyhydroxybutyrate (PHB) biosynthetic
CC pathway. It catalyses the reduction of acetoacetyl CoA to form
CC D(-)-beta-hydroxybutyryl-CoA, the substrate for PHB synthetase, and can
CC be used to control biopolymer syntheses and produce new biopolymers.
CC The sequence was deduced from the DNA sequence.
CC See also P94154-P94156.
SQ Sequence 246 AA;

P94157 Length: 246 February 11, 2000 15:48 Type: P Check: 2829 ..

1 MTQRIAYVTG GMGGIGTAIC QRLAKDGFRV VAGCGPNSPR REKWLEQQKA

51 LGFDFIASEG NVADWDSTKT AFDKVKSEVG EVDVLINNAG ITRDVVFRKM

101 TRADWDAVID TNLTSLFNVT KQVIDGMADR GWGRIVNISS VNGQKGQFGQ

151 TNYSTAKAGL HGFTMALAQE VATKGVTVNT VSPGYIATDM VKAIRQDVLD

201 KIVATIPVKR LGLPEEIASI CAWLSSEESG FSTGADFSLN GGLHMG

!!AA_SEQUENCE 1.0
ID P90460 standard; protein; 288 AA.
AC P90460;
DT 1-NOV-1989 (first entry)
DE Alpha-factor profibrolase from Southern Copperhead venom
KW Southern copperhead fibrolase; snake venom fibrolase;
KW alpha-factor profibrolase fusion construct; Crotalidus;
KW thrombolytic agents; plasmid pKS308.
OS Agkistrodon contortrix contortrix
PN EP-323722-A.
PD 12-JUL-1989.
PF 16-DEC-1988; 311924.
PR 18-DEC-1987; US-134981.
PA (CHIR) Chiron Corp.
PI Valenzuela P, Brake A, Randolph A;
DR WPI; 89-200623/28.
DR N-PSDB; N90243.
PT DNA encoding snake venom fibrolase
PT - for use as thrombolytic agent.
PS Disclosure; fig 8; 41pp; English.
CC Alpha-factor profibrolase fusion construct (see N90243) of
CC plasmid pKS308. Used as a thrombolytic agent to treat
CC vertebrates, esp. humans. Produced by recombinant techniques.
SQ Sequence 288 AA;

P90460 Length: 288 February 11, 2000 15:48 Type: P Check: 560 ..

1 MRFPSIFTAV LFAASSALAA PVNTTTEDET AQIPAEAVIG YLDLEGDFDV

51 AVLPFSNSTN NGLLFINTTI ASIAAKEEGV SLDKRQQRFP QRYVQLVIVA

101 DHRMNTKYNG DSDKIRQWVH QIVNTINEIY RPLNIQFTLV GLEIWSNQDL

151 ITVTSVSHDT LASFGNWRET DLLRRQRHDN AQLLTAIDFD GDTVGLAYVG

201 GMCQLKHSTG VIQDHSAINL LVALTMAHEL GHNLGMNHDG NQCHCGANSC

251 VMAAMLSDQP SKLFSDCSKK DYQTFLTVNN PQCILNKP

```
!!AA_SEQUENCE 1.0
ID   P81999 standard; protein; 25 AA.
AC   P81999;
DT   18-OCT-1990  (first entry)
DE   Binding peptide 2 capable of binding to rotavirus capsid protein VP6
KW   rotavirus strain S-A11; VP6 inner capsid protein; binding peptide;
KW   gastroenteritis; ss.
OS   Rotavirus strain S-A11.
PN   EP-259149-A.
PD   09-MAR-1988.
PF   02-SEP-1987; 307746.
PR   03-SEP-1986; US-903222.
PA   (UYSA-) Univ Saskatchewan.
PI   Sabara M I, Frenchick P J, Mullin-Ready K F;
DR   WPI; 88-065927/10.
PT   Inducing immune response with complex of epitope contg molecule -
PT   and rotavirus VP6 inner capsid protein as carrier, attached
PT   without chemical coupling agent
PS   Claim 9; Page 22; 30pp; English.
CC   This peptide is found to bind to rotavirus protein VP6. It is
CC   derived from rotavirus VP3 protein. Ability to bind is due to
CC   the spatial arrangement of a Cys and an Arg in the 3-D protein.
CC   VP6 is used as a carrier in a compsn for inducing immunological
CC   response in mammals to a selected epitope. The carrier protein is
CC   coupled to an epitope-bearing molecule via a binding sequence.
CC   (See also P81997 and P81999-P82000).
CC   Use of the VP6 inner capsid protein as carrier eliminates the need
CC   for coupling agents (protein-protein interaction used) and avoids
CC   adverse effects or antigenicity.
SQ   Sequence   25 AA;

 P81999  Length: 25  February 11, 2000 15:48  Type: P  Check: 4995  ..

       1  CNIAPASIVS RNIVYTRAQP NQDIA

!!AA_SEQUENCE 1.0
ID   P70477 standard; protein; 194 AA.
AC   P70477;
DT   05-APR-1991  (first entry)
DE   Sequence of human respiratory syncytial virus (HRSV) A2 strain
DE   22K protein.
KW   Vaccine.
OS   Human respiratory syncytial virus (HRSV).
PN   WO8704185-A.
PD   16-JUL-1987.
PF   23-DEC-1986; U02756.
PR   14-JAN-1986; US-818740.
PA   (UYNC-) UNIV OF N CAROLINA.
PA   (WERT/) WERTZ G W.
PA   (WERT/) WERTZ G W.
DR   WPI; 87-206300/29.
DR   N-PSDB; N70785.
PT   Vaccines for human respiratory virus - comprising proteins or
PT   fragment encoded by a DNA sequence coding for human respiratory
PT   syncytial virus proteins.
PS   Disclosure; Chart 14; 57pp; English.
CC   A novel plasmid which comprises a DNA sequence encoding this
CC   protein, and the protein itself, are claimed, for use as HRSV
CC   vaccines.  The vaccine can be administered to pregnant women or to
CC   women of child bearing age to stimulate maternal antibodies.
CC   Infants can also be vaccinated at 2-3 months of age.
SQ   Sequence   194 AA;

 P70477  Length: 194  February 11, 2000 15:48  Type: P  Check: 9828  ..

       1  MSRRNPCKFE IRGHCLNGKR CHFSHNYFEW PPHALLVRQN FMLNRILKSM

      51  DKSIDTLSEI SGAAELDRTE EYALGVVGVL ESYIGSINNI TKQSACVAMS

     101  KLLTELNSDD IKKLRDNEEL NSPKIRVYNT VISYIESNRK NNKQTIHLLK

     151  RLPADVLKKT IKNTLDIHKS ITINNPKEST VSDTNDHAKN NDTT

!!AA_SEQUENCE 1.0
ID   P70785 standard; protein; 194 AA.
AC   P70785;
DT   05-APR-1991  (first entry)
DE   Sequence encoding human respiratory syncytial virus (HRSV) A2 strain
DE   22K protein.
KW   Vaccine; ss.
OS   Human respiratory syncytial virus (HRSV).
PN   WO8704185-A.
PD   16-JUL-1987.
PF   23-DEC-1986; U02756.
PR   14-JAN-1986; US-818740.
PA   (UYNC-) UNIV OF N CAROLINA.
PA   (WERT/) WERTZ G W.
PA   (WERT/) WERTZ G W.
DR   WPI; 87-206300/29.
DR   P-PSDB; P70477.
PT   Vaccines for human respiratory virus - comprising proteins or
PT   fragment encoded by a DNA sequence coding for human respiratory
PT   syncytial virus proteins.
PS   Disclosure; Chart 14; 57pp; English.
CC   A novel plasmid which comprises a DNA sequence encoding this
CC   protein, and the protein itself, are claimed, for use as HRSV
CC   vaccines.  The vaccine can be administered to pregnant women or to
CC   women of child bearing age to stimulate maternal antibodies.
CC   Infants can also be vaccinated at 2-3 months of age.
SQ   Sequence   194 AA;

 P70785  Length: 194  February 11, 2000 15:48  Type: P  Check: 9828  ..

       1  MSRRNPCKFE IRGHCLNGKR CHFSHNYFEW PPHALLVRQN FMLNRILKSM

      51  DKSIDTLSEI SGAAELDRTE EYALGVVGVL ESYIGSINNI TKQSACVAMS

     101  KLLTELNSDD IKKLRDNEEL NSPKIRVYNT VISYIESNRK NNKQTIHLLK

     151  RLPADVLKKT IKNTLDIHKS ITINNPKEST VSDTNDHAKN NDTT

!!AA_SEQUENCE 1.0
ID   R24296 standard; Protein; 384 AA.
AC   R24296;
DT   20-NOV-1992  (first entry)
DE   Regulatory protein VanS involved in glycopeptide resistance.
KW   Glycopeptide antibiotic; vancomycin; teicoplanin; resistant;
KW   D-Ala-D-Ala ligase; peptidoglycan precursor; transposon;
KW   inverted repeats.
OS   Enterococcus faecium BM4147.
PN   WO9207942-A.
PD   14-MAY-1992.
PF   29-OCT-1991; F00855.
PR   31-OCT-1990; FR-013579.
PA   (INSP ) INST PASTEUR.
PI   Arthur M, Courvalin P, Dutka-malen S, Molinas C;
DR   WPI; 92-183677/22.
DR   N-PSDB; Q25178.
PT   Polypeptides involved in expression of glycopeptide antibiotic
PT   resistance - useful in diagnosing presence of Gram-positive
PT   enterococcal strains e.g. Enterococcus faecium and E gallinarum
PS   Claim 5; Fig 8; 163pp; French.
CC   VanS is coded for by a 7.3kb HindIII-EcoRI fragment of
CC   the plasmid pIP816. (The plasmid contains resistance genes from
CC   Enterococcus faecium BM 4147 and was described in New England J.
CC   Med., 319:157-161). The VanS protein is part of a two component
CC   regulatory system for controlling the amount of expression of the
CC   vancomycin-resistance proteins VanH, VanA and VanX. The other
CC   component of the regulatory system is the VanR transcription
CC   regulator. VanS protein can modify phosphorylation of VanR in
CC   response to vancomycin and thereby stimulates VanR-dependent
CC   transcription. See also Q25178-Q25183.
SQ   Sequence   384 AA;

 R24296  Length: 384  February 11, 2000 15:48  Type: P  Check: 3600  ..
```

```
       1  LVIKLKNKKN DYSKLERKLY MYIVAIVVVA IVFVLYIRSM IRGKLGDWIL

      51  SILENKYDLN HLDAMKLYQY SIRNNIDIFI YVAIVISILI LCRVMLSKFA

     101  KYFDEINTGI DVLIQNEDKQ IELSAEMDVM EQKLNTLKRT LEKREQDAKL

     151  AEQRKNDVVM YLAHDIKTPL TSIIGYLSLL DEAPDMPVDQ KAKYVHITLD

     201  KAYRLEQLID EFFEITRYNL QTITLTKTHI DLYYMLVQMT DEFYPQLSAH

     251  GKQAVIHAPE DLTVSGDPDK LARVFNNILK NAAAYSEDNS IIDITAGLSG

     301  DVVSIEFKNT GSIPKDKLAA IFEKFYRLDN ARSSDTGGAG LGLAIAKEII

     351  VQHGGQIYAE SNDNYTTFRV ELPAMPDLVD KRRS
```

```
!!AA_SEQUENCE 1.0
ID   R10995 standard; Protein; 398 AA.
AC   R10995;
DT   13-MAY-1991  (first entry)
DE   Xenopus Bone Morphogenetic Factor BMP-2A.
KW   BMF; osteoporosis; fracture; cartilage.
OS   Xenopus laevis.
PN   EP-416578-A.
PD   13-MAR-1991.
PF   05-SEP-1990; 117079.
PR   06-SEP-1989; JP-229250.
PR   20-JUL-1990; JP-190774.
PA   (TAKE ) TAKEDA CHEMICAL IND KK.
PA   (SCIT-) SCITECH RESEARCH CO.
PI   Murakami K, Ueno N, Kato Y;
DR   WPI; 91-075112/11.
DR   N-PSDB; Q10895.
PT   Xenopus laevis bone morphogenetic protein and DNA encoding it -
PT   used in therapy of fracture or osteoporosis
PS   Claim 2; Fig 4; 28pp; English.
CC   A Xenopus laevis unfertilised egg cDNA library in lambda gt10 was
CC   screened with a PstI-HindIII fragment of X.laevis Xar14 chromosomal
CC   DNA. Three clones were isolated, including clone Xbr22 which was
CC   found to encode a protein having homology with X.laevis BMP-2A. They
CC   were subcloned in pUC19 and used to transform competent E.coli HB101
CC   cells. Transformant E.coli HB101/pXbr22 coding for the BMP-2A was
CC   sequenced and the amino acid sequence of BMP-2A deduced from it.
CC   See also Q10890-4 and Q10896-7.
SQ   Sequence   398 AA;

 R10995  Length: 398  February 11, 2000 15:48  Type: P  Check: 1620  ..

       1  MVAGIHSLLL LQFYQILLSG CTGLVPEEGK RKYSESTRSS PQQSQQVLDQ

      51  FELRLLNMFG LKRRPTPGKN VVIPPYMLDL YHLHSAQLAD DQGSSEVDYH

     101  MERAASRANT VRSFHHEESM EEIPESGEKT IQRFFFNLSS IPDEELVTSS

     151  ELRIFREQVQ EPFKTDGSKL HRINIYDIVK PAAAASRGPV VRLLDTRLIH

     201  HNESKWESFD VTPAITRWIA HKQPNHGFVV EVTHLDNDTN VPKRHVRISR

     251  SLTLDKGHWP RIRPLLVTFS HDGKGHALHK RQKRQARHKQ RKRLKSSCRR

     301  HPLYVDFSDV GWNDWIVAPP GYHAFYCHGE CPFPLADHLN STNHAIVQTL

     351  VNSVNTNIPK ACCVPTELSA ISMLYLDENE KVVLKNYQDM VVEGCGCR
```

```
!!AA_SEQUENCE 1.0
ID   R24190 standard; Protein; 186 AA.
AC   R24190;
DT   24-NOV-1992  (first entry)
DE   Bovine RSV strain A 51908 M2 protein.
KW   Bovine respiratory syncytial virus; vaccine; diagnosis; antibodies;
KW   M2 gene; BRSV.
OS   Bovine respiratory syncytial virus strain A 51908.
PN   WO9207940-A.
PD   14-MAY-1992.
PF   04-NOV-1991; U08177.
PR   05-NOV-1990; US-608937.
PA   (SAMA/) SAMAL S K.
PI   Samal SK;
DR   WPI; 92-183675/22.
DR   N-PSDB; Q25033.
PT   Bovine respiratory syncytial virus genes - used in the prodn. of
PT   agents for use in detection and as vaccines for BRSV infection.
PS   Claim 32; Page 57; 74pp; English.
CC   This is the sequence of bovine respiratory syncytial virus (BRSV)
CC   strain A51908 M2 protein. It can be used in the detection of BRSV
CC   antibodies and in vaccines to prevent infection. It can also be
CC   used for the production of BRSV protein antibodies.
CC   See also R24184-R24191 and R25310.
SQ   Sequence   186 AA;

 R24190  Length: 186  February 11, 2000 15:48  Type: P  Check: 5872  ..

       1  MSRRNPCKYE IRGHCLNGKK CHFSHNYFEW APHALLVRQN FMLNKILKSM

      51  DRNNDTLSEI SGAAELDRTE EYALGVIGVL ESYLSSINNI TKQSACVAMS

     101  KLLAEINNDD IKRLRNKEVP TSPKIRIYNT VISYIDSNKR NTKQTIHLLK

     151  RLPADVLKKT IKNTIDIHNE INGNNQGDIN VDEQNE
```

```
!!AA_SEQUENCE 1.0
ID   P81092 standard; protein; 95 AA.
AC   P81092;
DT   21-MAR-1991  (first entry)
DE   Sequence of rhinovirus HRV2 viral protein P2B.
KW   Passive immunity; diagnosis; therapy; ss.
OS   Rhinovirus.
PN   EP-261403-A.
PD   30-MAR-1988.
PF   20-AUG-1987; 112104.
PR   17-JAN-1987; DE-701301.
PA   (BOEH) Boehringer Ingelheim.
PI   Duechler M, Skern T, Sommergruber W, Neubauer C, Grundler P, Blaas,
PI   Kuchler E, Frasei L, Zorn M;
DR   WPI; 88-085735/13.
PT   New DNA corresponding to viral RNA of rhino-virus HRV89 - useful for
PT   prodn. of polypeptide(s) for stimulating immune system against HRV
PT   89.
PS   Example; Fig 20; 66pp; German.
CC   DNA molecules corresponding to all or part of the RNA of rhinovirus
CC   strain HRV89 (Fig 4, N81390) is claimed, esp. the portion encoding
CC   the viral proteins VP1-VP4, P2A-P2C, P3A-P3C. Also claimed are the
CC   polypeptides encoded by any of these DNA molecules. The polypeptides
CC   are used for stimulating a protective immune response and for
CC   blocking cellular receptors. Ab are useful  for assay and for
CC   purificn. of the corresp. antigen, and can also be used for the
CC   therapeutic and diagnostic applications.
SQ   Sequence   95 AA;

 P81092  Length: 95  February 11, 2000 15:48  Type: P  Check: 3060  ..

       1  GVTDYIHMLG EAFGNGFVDS VKEHIHAINP VGNISKKIIK WMLRIISAMV

      51  IIIRNSSDPQ TILATLTLIG CSGSPWRFLK EKFCKWTQLN YIHKE
```

```
!!AA_SEQUENCE 1.0
ID   P81093 standard; protein; 323 AA.
AC   P81093;
DT   21-MAR-1991  (first entry)
DE   Sequence of rhinovirus HRV2 viral protein P2C.
KW   Passive immunity; diagnosis; therapy; ss.
OS   Rhinovirus.
PN   EP-261403-A.
PD   30-MAR-1988.
PF   20-AUG-1987; 112104.
```

PR 17-JAN-1987; DE-701301.
PA (BOEH) Boehringer Ingelheim.
PI Duechler M, Skern T, Sommergruber W, Neubauer C, Grundler P, Blaas,
PI Kuchler E, Frasei L, Zorn M;
DR WPI; 88-085735/13.
PT New DNA corresponding to viral RNA of rhino-virus HRV89 - useful for
PT prodn. of polypeptide(s) for stimulating immune system against HRV
PT 89.
PS Example; Fig 20; 66pp; German.
CC DNA molecules corresponding to all or part of the RNA of rhinovirus
CC strain HRV89 (Fig 4, N81390) is claimed, esp. the portion encoding
CC the viral proteins VP1-VP4, P2A-P2C, P3A-P3C. Also claimed are the
CC polypeptides encoded by any of these DNA molecules. The polypeptides
CC are used for stimulating a protective immune response and for
CC blocking cellular receptors. Ab are useful for assay and for
CC purificn. of the corresp. antigen, and can also be used for the
CC therapeutic and diagnostic applications.
SQ Sequence 323 AA;

P81093 Length: 323 February 11, 2000 15:48 Type: P Check: 5281 ..

1 SDSWLKKFTE ACNAARGLEW IGNKISKFIE WMKSMLPQAQ LKVKYLNELK

51 KLNLYEKQVE SLRAVDMKTQ EKIKMEIDTL HDLSRKFLPL YASEAKRIKT

101 LYIKCDNIIK QKKRCEPVAI VIHGPPGAGK SITTNFLAKM ITNDSDIYSL

151 PPDPKYFDGI EDQQSVVIMD DIMQNPAGDD MTLFCQMVSS VTFIPPMADL

201 PDKGKAFDSR FVLCSTNHSL LTPPTITSLP AMNRRFFLDL DIIVHDNFKD

251 PQGKLNVAAA FRPCDVDNRI GNARCCPFVC GKAVSFKDRN SCNKYSLAQV

301 YNIMIEEDRR RRQVVDVMTA IFQ

!!AA_SEQUENCE 1.0
ID R10974 standard; Protein; 246 AA.
AC R10974;
DT 17-APR-1991 (first entry)
DE Acetoacetyl CoA reductase enzyme.
KW Polyester biopolymers; polyhydroxybutyrate; polyhydroxy alkanoate;
KW beta-ketothiolase; acetoacetyl CoA reductase.
OS Alcaligenes eutrophus.
PN WO9100917-A.
PD 24-JAN-1991.
PF 10-JUL-1990; U03851.
PR 10-JUL-1989; US-378155.
PA (MASI) MASSACHUSETTS INST TECH.
PI Peoples OP, Sinskey AJ;
DR WPI; 91-051341/07.
DR N-PSDB; Q10501.
PT Construction and modification of polyester bio:polymers - by
PT introduction of poly-hydroxy-butyrate and -alkanoate genes into
PT bacteria or plants
PS Disclosure; fig 3; 64pp; English.
CC This Alcaligenes eutrophus acetyl CoA reductase is an enzyme
CC which is essential to the biosynthesis of polyhydroxyalkanoate
CC (PHA). The gene encoding this is contained in plasmid clone,
CC pAeT3, downstream from the thiolase gene. The use of recombinant
CC methods for producing such enzymes, required for polyester bio-
CC polymer synthesis, allows for the control and modification of the
CC synthesis process.
CC See also Q10499-500 and Q10502-03.
SQ Sequence 246 AA;

R10974 Length: 246 February 11, 2000 15:48 Type: P Check: 3836 ..

1 MTQRIAYVTG GMGGIGTAIC QRLAKDGFRV VAGCGPNSPR REKWLEQQKA

51 LGYDFIASEG NVADWDSTKT AFDKVKSEVG EVDVLINNAG ITRDVVFRKM

101 TRADWDAVID TNLTSLFNVT KQVIDGMADR GWGRIVNISS VNGQKGQFGQ

151 TNYSTAKAGL HGFTMALAQE VATKGVTVNT VSPGYIATDM VKAIRQDVLD

201 KIVATIPVKR LGLPEEIASI CAWLSSEESG FSTGADFSLN GGLHMG

!!AA_SEQUENCE 1.0
ID R12468 standard; Protein; 248 AA.
AC R12468;
DT 29-AUG-1991 (first entry)
DE Luffa cylindrica bioactive protein.
KW Protein synthesis inhibition.
OS Luffa cylindrica.
PN J03109398-A.
PD 09-MAY-1991.
PF 22-SEP-1989; 247410.
PR 22-SEP-1989; JP-247410.
PA (SUMO) SUMITOMO CHEM IND KK.
DR WPI; 91-181458/25.
PT New protein of Luffa cylindrica, extracted form seeds - is used for
PT inhibiting protein synthesis in viral proliferation.
PS Claim 1; Fig 1; 6pp; Japanese.
CC The protein is prepd. by extraction from seeds of L. cylindrica.
CC It has protein synthesis inhibiting activity and is useful for
CC inhibiting viral growth.
CC See also J0112999.
SQ Sequence 248 AA;

R12468 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..

1 DVRFSLSGSS STSYSKFIGD LRKALPSNGT VYNLTILLSS ASGASRYTLM

51 TLSNYDGKAI TVAVDVSQLY IMGYLVNSTS YFFNESDAKL ASQYVFKGST

101 IVTLPYSGNY EKLQTAAGKI REKIPLGFPA LDSALTTIFH YDSTAAAAAF

151 LVILQTTAEA SRFKYIEGQI IERISKNQVP SLATISLENS LWSALSKQIQ

201 LAQTNNGTFK TPVVITDDKG QRVEITNVTS KVVTKNIQLL LNYKQNVA

!!AA_SEQUENCE 1.0
ID R13252 standard; Protein; 247 AA.
AC R13252;
DT 11-OCT-1991 (first entry)
DE Murine Cytotoxic Cell Protease-1.
KW mouse; CCP-1 inhibitor; cytotoxic T-lymphocytes; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT peptide 1. .20
FT /label= signal_peptide
FT protein 21. .247
FT /label= CCP-1
PN WO9110685-A.
PD 25-JUL-1991.
PF 17-JAN-1991; U00340.
PR 19-JAN-1990; US-467880.
PA (SERA-) SERAGEN INC.
PI Bleackley RC, Lobe CG, Paetkau VH, James MN, Murphy M;
DR WPI; 91-237989/32.
DR N-PSDB; Q12862.
PT DNA vectors, and inhibitors of cytotoxic cell protease - for
PT treatment of auto-immune diseases e.g. pernicious anaemia,
PT rheumatoid arthritis, allo-graft rejection etc.
PS Claim 5; Fig 3; 62pp; English.
CC The CCP-1 coding sequence was isolated from the cytotoxic T-cell
CC line MTL 2.8.2 generated from CBA/J mice. The amino acid sequence
CC of CCP-1 was predicted from the cDNA sequence. The structure of the
CC protein was used to design peptides which competitively inhibit the
CC protease. See also Q12863-6 and R13254-R13262.
SQ Sequence 247 AA;

R13252 Length: 247 February 11, 2000 15:48 Type: P Check: 528 ..

1 MKILLLLLTL SLASRTKAGE IIGGHEVKPH SRPYMALLSI KDQQPEAICG

```
      51  GFLIREDFVL TAAHCEGSII NVTLGAHNIK EQEKTQQVIP MVKSIPHPDY

     101  NPKTFSNDIM LLKLKSKAKR TRAVRPLNLP RRNVNVKPGD VCYVAGWGRM

     151  APIGKYSNTL QEVELTVQKD RECESYFKNR YNKTNQICAG DPKTKRASFR

     201  GDSGGPLVCK KVAAGIVSYG YKDGSPPRAF TKVSSFLSWI KKTMKSS

!!AA_SEQUENCE 1.0
ID   P60645 standard; protein; 292 AA.
AC   P60645;
DT   10-MAR-1993  (revised)
DT   01-JAN-1980  (first entry)
DE   Mouse kidney cell Band 3-like anion transport protein.
KW   Band 3; anion transport protein; glycoprotein; cystic fibrosis;
KW   disease diagnosis; erythrocyte; ss.
OS   Mus musculus.
PN   WO8605203-A.
PD   12-SEP-1986.
PF   06-MAR-1986; U00450.
PR   07-MAR-1985; US-709183.
PA   (WHIT-) WHITEHEAD INST BIOM.
PI   Alper SI, Kopito RR;
DR   WPI; 86-252231/38.
DR   NP-PSDB; N60551.
PT   DNA coding for anion transport protein - esp. for detection of
PT   genetic defects
PS   Disclosure; Fig. 9; 49pp; English.
CC   Band 3 is the major glycoprotein of mammalian erythrocytes and
CC   functions as an anion antiport in that it mediates the 1-for-1
CC   exchange of chloride (Cl-) and bicarbonate (HCO3-) across the
CC   erythrocyte plasma membrane.  The anion transport protein may be
CC   used in the detection of genetic defects resulting in abnormal
CC   transport of ions across cell membranes, e.g. as in cystic fibrosis.
CC   This protein is specifically from mouse kidney cells.  See also
CC   N60550 and P60644.
SQ   Sequence   292 AA;

 P60645  Length: 292  February 11, 2000 15:48  Type: P  Check: 2381  ..

       1  VDYSIEDTYT QKLSVPSGFS VTAPDKRGWV INPLGEKTPF PVWMMVASLL

      51  PAVLVFILIF METQITTLII SKKERMLQKG SGFHLDLLLI VAMGGICALF

     101  GLPWLAAATV RSVTHANALT VMSKAVAPGD KPKIQEVKEQ RVTGLLVALL

     151  VGLSMVIGDL LRQIPLAVLF GIFLYMGVTS LNGIQFYERL HLLLMPPKHQ

     201  PDVTYVKKVR TMRMHLFTAL QLLCLALLWA VMSTAASLAF PFILILTVPL

     251  RMVVLTRIFT EREMKCLDAN EAEPVFDECE GVDEYNEMPM PV

!!AA_SEQUENCE 1.0
ID   R14247 standard; Protein; 149 AA.
AC   R14247;
DT   03-JAN-1992  (first entry)
DE   HCV-T (6110-6557) encoded epitope.
KW   Epitope; PCR; diagnosis.
OS   Hepatitis C virus.
PN   WO9114779-A.
PD   03-OCT-1991.
PF   28-MAR-1991; J00405.
PR   28-MAR-1990; JP-080185.
PR   13-JUN-1990; JP-154230.
PR   14-JUN-1990; JP-153979.
PR   09-NOV-1990; JP-305795.
PA   (MITK ) MITSUI TOATSU CHEM INC.
PI   Takada T, Enomoto N, Date T, Nakao T;
DR   WPI; 91-310579/42.
DR   N-PSDB; Q14076.
PT   New nucleotide sequences encoding HCV epitope(s) - for diagnosis
PT   of hepatitis C virus infection via polymerase chain reaction
PS   Disclosure; Fig 1(A)-(B); 79pp; Japanese.
CC   The nucleotide sequences represented in Q14076-86 and Q14767-71
CC   encode epitopes from structural, non-structural and 5' untranslated
CC   domains of hepatitis C virus.  The sequences are used for accurate
CC   and simple diagnosis and typing of HCV infection, using PCR
CC   amplification techniques.
SQ   Sequence   149 AA;

 R14247  Length: 149  February 11, 2000 15:48  Type: P  Check: 3078  ..

       1  VLKESKAKAS TVKAKLLSVE EACKLTPPHS ARSKFGYGAK DVRSLSSRAI

      51  NHIRSVWKDL LEDTETPIDT TIMAKNEVFC VQPEKGGRKP ARLIVFPDLG

     101  VRVCEKMALY DVVSTLPQAV MGSSYGFQYS PGQRVEFLVN AWKSKKCPM

!!AA_SEQUENCE 1.0
ID   R20004 standard; peptide; 25 AA.
AC   R20004;
DT   25-MAR-1992  (first entry)
DE   VP6 inner capsid epitope, 84 TS, Peptide B.
KW   Rotavirus; vaccine; gastroenteritis; virus protein 6; hapten;
KW   immunogen; conjugate; carrier.
OS   Synthetic.
PN   US5071651-A.
PD   10-DEC-1991.
PF   05-MAR-1990; 489790.
PR   03-SEP-1986; US-903222.
PR   02-SEP-1987; US-092120.
PR   05-MAR-1990; US-489790.
PA   (UYSA-) UNIV OF SASKATCHEWAN.
PI   Sabara MI, Frenchick PJ, Mullin-Ready KF;
DR   WPI; 92-006804/01.
PT   New immunological carrier complex - contains VP6 polypeptide from
PT   rota:virus as the carrier mol.
PS   Claim 9b; Page 15; 23pp; English.
CC   The peptide binds to purified VP6 from bovine rotavirus strain C486.
CC   It was derived from the trypsin cleavage site of VP3 outer capsid
CC   protein and corresponds to amino acids 231-254.  The cysteine at
CC   position 1 was added to facilitate coupling to a carrier protein.
CC   The peptide has an apparent mol. wt. of 2,734.  The peptide can be
CC   used in vaccines against rotavirus, or to construct immunoconjugates
CC   with VP6 particles as carriers for haptens.
CC   See also R20003, and R21043-48, and R21071-73.
SQ   Sequence   25 AA;

 R20004  Length: 25  February 11, 2000 15:48  Type: P  Check: 4995  ..

       1  CNIAPASIVS RNIVYTRAQP NQDIA

!!AA_SEQUENCE 1.0
ID   R21409 standard; Protein; 482 AA.
AC   R21409;
DT   01-APR-1992  (first entry)
DE   NADH dehydrogenase subunit 2.
KW   Pneumonia; assay; AIDS; immunosuppressed.
OS   Pneumocystis carinii.
PN   WO9119005-A.
PD   12-DEC-1991.
PF   31-MAY-1991; G00869.
PR   01-JUN-1990; GB-012196.
PA   (ISIS-) ISIS INNOVATION LTD.
PI   Wakefield AE, Hopkin JM, Moxon ER;
DR   WPI; 92-007487/01.
DR   N-PSDB; Q20065.
PT   New DNA sequences which act as oligo:nucleotide primers - for
PT   assaying DNA sample from respiratory secretion of a patient
PT   infected with P carinii
PS   Claim 3; Fig 3; 42pp; English.
CC   The amino acid sequence is that of P. carinii NADH dehydrogenase
CC   subunit 2 which was translated from DNA from plasmid pAZ112. See
CC   also R21410-R21413 and R20056.
SQ   Sequence   482 AA;
```

R21409 Length: 482 February 11, 2000 15:48 Type: P Check: 617 ..

```
      1  MLLSSIISQL IAIALSSSHW NLVLLSRISI ISLIYSIILT YNVYYVEIIG

     51  LGLGIYNGFL QVTSLTQFVD IFIFLLGILI LGITGFYHVD KDNSRELTSL

    101  YDFKQHLEYP VLSLFLLFGS QILVSSLNVI TFYLSLELQS FSLYILSSLR

    151  SSKQGLKYFL LGALSSCFIL LGFGLVYSYT GITSLESLAI FSKVNLNIYM

    201  QISLLICVLG ILFKIGIVPF HQWAIDVYDG VPTIITTWLT TLTKISLLIF

    251  LMEFIYHHSS ENWTTILMLL SVLSVIVGSL LGLSQSRIKR LLIYSMVSHV

    301  GFLMLSLSIM TEKSLEAFLF YLVQYSITNL NVFLILIAMG YFYKNPDSED

    351  SPIIYINSLR GLVRVQPLLS ICLAISLLSL GGIPPFIGFF GKLNILYSTI

    401  TQGYLFISIL LVLASVLSIS YYLKVVQLLF VGESSLSFRN IQISTYLSTL

    451  IGVLTLMIAM FLVNPDFILQ LINITICKYF IL
```

!!AA_SEQUENCE 1.0
ID R23004 standard; Protein; 297 AA.
AC R23004;
DT 29-OCT-1992 (first entry)
DE Protein transcribed from the mba sequence of plasmid pKDSC50.
KW Virulent; Salmonella choleraesuis; mouse bacteremia; Salmonella.
OS Salmonella choleraesuis.
PN J04079885-A.
PD 13-MAR-1992.
PF 24-JUL-1990; 194069.
PR 24-JUL-1990; JP-194069.
PA (KIBU-) KIBUN KK.
PA (KITA) KITASATO RES INST.
DR WPI; 92-137918/17.
DR N-PSDB; Q24046.
PT New virulence plasmid pKDSC50 - provides DNA fragment useful as
PT probe for detecting Salmonella
PS Disclosure; Fig 2; 9pp; Japanese.
CC The 32 kD protein may be transcribed from the first reading frame of
CC the "mba" fragment of virulent plasmid pKDSC50 from nucleotide 1159-
CC 2052. Three other proteins may also be transcribed from the Salmonella
CC "mba" fragment, which is capable of causing mouse bacteremia.
CC See also R23005-7.
SQ Sequence 297 AA;

R23004 Length: 297 February 11, 2000 15:48 Type: P Check: 7059 ..

```
      1  MDFLINKKLK IFITLMETGS FSIATSVLYI TRTPLSRVIS DLERELKQRL

     51  FIRKNGTLIP TEFAQTIYRK VKSHYIFLHA LEQEIGPTGK TKQLEIIFDE

    101  IYPESLKNLI ISALTISGQK TNIMRRAVNS QIIEELCQTN NCIVISARNY

    151  FHRESLVCRT SVEGGVMLFI PKKFFLCGKP DINRLAGTPV LFHEGAKNFN

    201  LDTIYHFFEQ TLGITNPAFS FDNVDLFSSL YRLQQGLAML LIPVRVCRAL

    251  GLSTDHALHI KGVALCTSLY YPTKKRETPD YRKAIKLIQQ ELKQSTF
```

!!AA_SEQUENCE 1.0
ID R22695 standard; Protein; 286 AA.
AC R22695;
DT 10-NOV-1992 (first entry)
DE blaC31 mutation (5).
KW Beta-lactamase; BspMI excision linker; human growth hormone;
KW antibiotic; transformants.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_difference 194. .198
FT /label= MC-16
PN WO9207090-A.
PD 30-APR-1992.
PF 21-OCT-1991; U07506.
PR 22-OCT-1990; US-602158.
PA (GETH) GENENTECH INC.
PI Botstein D, Palzkill T;
DR WPI; 92-167168/20.
DR N-PSDB; Q24133.
PT Novel DNA modification method and prodn. of modified DNA
PT libraries - used to determine effect of modification on
PT interaction of encoded product, e.g. enzyme with target
PS Disclosure; Fig 10; 98pp; English.
CC The sequence given is the beta-lactamase (bla) gene product which
CC has been modified using the BspMI excision linker MC-16 which is
CC one of a group of linkers designated blaC31.
CC Within the scope of the invention, this sequence was used to
CC illustrate a novel method of modifying DNA. DNA libraries
CC expressing these modified polypeptides can be used to analyse the
CC relationship between structure and function. Modified beta-lactamse
CC libraries can be used to determine the susceptibility of a particular
CC antibiotic to hydrolysis ie. neutralization by mutant enzymes. Since
CC all possible combinations of mutations were represented within the scope
CC of the invention, a meaningful sample could be achieved. The method of
CC the invention allows the neasurement of successful transformants which
CC contain a mutation linker, to determine the number of different colonies
CC present in the library. This can then be used to calculate the
CC probabilities that the most and the least common codon combinations are
CC present in the library. The libraries can also be used to produce sets
CC of modified polypeptides such as hormones, eg. human growth hormone, in
CC order to detect and anlyse structure/function relationships.
SQ Sequence 286 AA;

R22695 Length: 286 February 11, 2000 15:48 Type: P Check: 4919 ..

```
      1  MSIQHFRVAL IPFFAAFCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN

     51  SGKILESFRP EERFPMMSTF KVLLCGAVLS RVDAGQEQLG RRIHYSQNDF

    101  VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTTI GGPKELTAFL

    151  HNMGDHVTRL DRCEPELNEA IPNDERDTTM PAAMATTLRK LLTASSVTLA

    201  SRQQLRDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS RGIIAALGPD

    251  GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW
```

!!AA_SEQUENCE 1.0
ID R25601 standard; Protein; 85 AA.
AC R25601;
DT 18-JAN-1993 (first entry)
DE NANBH virus antigen.
KW Antibody; acute phase; chronic phase; vaccine.
OS Non-A, non-B hepatitis virus.
PN J04169193-A.
PD 17-JUN-1992.
PF 09-OCT-1990; 272980.
PR 09-OCT-1990; JP-272980.
PA (YAMA) YAMANOUCHI PHARM CO LTD.
DR WPI; 92-253396/31.
DR N-PSDB; Q26616.
PT Antigenic polypeptide of non-A, non-B hepatitis virus - for
PT diagnosing non-A hepatitis virus infections, also used as vaccine
PS Claim 7; Page 2; 10pp; Japanese.
CC The sequence given is an antigenic polypeptide of non-A, non-B
CC hepatitis (NANBH) virus. This polypeptide is useful for the detection
CC of NANBH virus antibodies. This antigen can detect the virus
CC antibodies in both the acute phase and chronic phase. The detection of
CC NANBH virus genome using this antigen is very sensitive. Polyclonal
CC and monoclonal antibodies for the detection of the virus antigen and
CC the medical treatment of NANBH, can be prepared using the antigenic
CC peptide. This peptide can be used in the production of a vaccine.
SQ Sequence 85 AA;

R25601 Length: 85 February 11, 2000 15:48 Type: P Check: 5700 ..

```
      1  KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC KLTPPHSAKS

     51  KFGYGAKDVR NLSSKAVNHI RSVWKDLLED TETPL

!!AA_SEQUENCE 1.0
ID   R25869 standard; Protein; 239 AA.
AC   R25869;
DT   21-JAN-1993  (first entry)
DE   HCV polypeptide 16.
KW   Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
OS   Hepatitis C virus.
PN   J04179482-A.
PD   26-JUN-1992.
PF   11-NOV-1990; 304417.
PR   11-NOV-1990; JP-304417.
PA   (TOKU ) TOKUYAMA SODA KK.
DR   WPI; 92-263663/32.
DR   N-PSDB; Q26996.
PT   Hepatitis C virus antigen expressed as recombinant in E.coli -
PT   useful for diagnosis of hepatitis C virus infection
PS   Claim 1; Page 5; 66pp; Japanese.
CC   The sequences given in R25854-74 are hepatitis C virus protiens.
CC   The genes encoding these proteins can each be used to prepare
CC   recombinant vectors by ligating the gene of interest in to a vector
CC   to be expressed in E. coli.  These polypeptides are useful as
CC   diagnostic reagents for type C hepatitis and they may be produced
CC   efficiently by recombinant methods.
SQ   Sequence   239 AA;

 R25869  Length: 239  February 11, 2000 15:48  Type: P  Check: 3783  ..

      1  VCCSMSYTWT GALITPCAAE ESKLPINPLS NSLLRHHSMV YSTTSRSASL

     51  RQKKVTFDRL QVLDDHYRDV LKEMKAKAST VKARLLSIEE ACKLTPPHSA

    101  KSKFGYGAKD VRSLSSRAVN HIRSVWEDLL EDTETPIDTT IMAKNEVFCV

    151  QPEKGGRKPA RLIVFPDLGV RVCEKMALYD VVSTLPQAVM GPSYGFQYSP

    201  GQRVEFLVNT WKSKKCPMGF SYDTRCFDST VTENDIRTE

!!AA_SEQUENCE 1.0
ID   R25891 standard; Protein; 253 AA.
AC   R25891;
DT   21-JAN-1993  (first entry)
DE   HK16.
KW   Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
OS   Hepatitis C virus.
FH   Key             Location/Qualifiers
FT   protein         13. .251
FT                   /*tag=  a
FT                   /note= "Sequence R25869"
PN   J04179482-A.
PD   26-JUN-1992.
PF   11-NOV-1990; 304417.
PR   11-NOV-1990; JP-304417.
PA   (TOKU ) TOKUYAMA SODA KK.
DR   WPI; 92-263663/32.
DR   N-PSDB; Q27018.
PT   Hepatitis C virus antigen expressed as recombinant in E.coli -
PT   useful for diagnosis of hepatitis C virus infection
PS   Disclosure; Fig 17; 66pp; Japanese.
CC   The sequences given in R25876-95 are encoded by the claimed hepatitis
CC   C virus genes of the invention which have been inserted into an E.
CC   coli vector.  These polypeptides are useful as diagnostic reagents
CC   for type C hepatitis and they may be produced efficiently by
CC   recombinant DNA techniques.
SQ   Sequence   253 AA;

 R25891  Length: 253  February 11, 2000 15:48  Type: P  Check: 4644  ..

      1  MITPSLHACM QVVCCSMSYT WTGALITPCA AEESKLPINP LSNSLLRHHS

     51  MVYSTTSRSA SLRQKKVTFD RLQVLDDHYR DVLKEMKAKA STVKARLLSI

    101  EEACKLTPPH SAKSKFGYGA KDVRSLSSRA VNHIRSVWED LLEDTETPID

    151  TTIMAKNEVF CVQPEKGGRK PARLIVFPDL GVRVCEKMAL YDVVSTLPQA

    201  VMGPSYGFQY SPGQRVEFLV NTWKSKKCPM GFSYDTRCFD STVTENDIRT

    251  EGL

!!AA_SEQUENCE 1.0
ID   R25303 standard; Protein; 194 AA.
AC   R25303;
DT   03-MAR-1993  (first entry)
DE   HRSV 22K protein.
KW   Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
KW   major capsid protein; N.
OS   Human respiratory syncytial virus strain A2.
PN   US5149650-A.
PD   22-SEP-1992.
PF   14-JAN-1986; 818740.
PR   14-JAN-1986; US-818740.
PR   13-JUL-1988; US-218737.
PA   (UYNC-) UNIV NORTH CAROLINA.
PI   Collins PL, Wertz GW;
DR   WPI; 92-340247/41.
DR   N-PSDB; Q29624.
PT   Vaccines for human respiratory virus - include structural genes
PT   coding for native structural viral proteins and immunogenic
PT   fragments
PS   Disclosure; Page 19; 21pp; English.
CC   The sequences of mRNA encoding HRSV structural proteins are given in
CC   Q29622-26.  The proteins are F, G, 22K, 9.5K and major capsid
CC   protein N.  The sequences and encoded proteins are useful for
CC   preparing vaccines against HRSV.  The vaccines can be used to confer
CC   immunity against respiratory tract infections on human subjects.
SQ   Sequence   194 AA;

 R25303  Length: 194  February 11, 2000 15:48  Type: P  Check: 9668  ..

      1  MSRTNPCKFE IRGHCLNGKR CHFSHNYFEW PPHALLVRQN FMLNRILKSM

     51  DKSIDTLSEI SGAAELDRTE EYALGVVGVL ESYIGSINNI TKQSACVAIS

    101  KLLTELNSDD IKKLRDNEEL NSPKIRVYNT VISYIESNRK NNKQTIHLLK

    151  RLPADVLKKT IKNTLDIHKS ITINNPKEST VSDTNDHAKN NDTT

!!AA_SEQUENCE 1.0
ID   R29874 standard; Protein; 171 AA.
AC   R29874;
DT   26-APR-1993  (first entry)
DE   HCV NS4-NS5 peptide N29-1, N29-2, N29-3.
KW   Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW   transcriptase; cDNA; primer; allele; core; region; upstream;
KW   hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW   determinant; antiserum.
OS   Hepatitis C virus
FH   Key             Location/Qualifiers
FT   misc_difference 61
FT                   /label= Glu, Lys
FT   misc_difference 155
FT                   /label= Ala, Ser
PN   EP-518313-A.
PD   16-DEC-1992.
PF   11-JUN-1992; 109812.
PR   11-JUN-1991; JP-139268.
PR   12-JUL-1991; JP-172794.
PR   07-OCT-1991; JP-287008.
PR   16-DEC-1991; JP-332329.
PR   20-APR-1992; JP-099957.
PA   (MITU ) MITSUBISHI KASEI CORP.
```

PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI; 92-417213/51.
DR N-PSDB; Q32505.
PT New hepatitis C virus gene and its encoded protein - used for
PT diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure; Page 197-98; 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequencess were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in Q32565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 171 AA;

R29874 Length: 171 February 11, 2000 15:48 Type: P Check: 9846 ..

1 YRDVLKEMKA KASTVKAKLL SVEEACKLTP PHSARSKFGY GAKDVRSLSS

51 KAVNHIRSVW XDLLEDTDTP IDTTIMAKNE VFCVQPEKGG RKPARLIVFP

101 DLGVRVCEKM ALYDVVSTLP QAVMGSSYGF QYSPGQRVEF LVNAWKSKKS

151 PMGFXYDTRC FDSTVTENDI R

!!AA_SEQUENCE 1.0
ID R29879 standard; Protein; 285 AA.
AC R29879;
DT 26-APR-1993 (first entry)
DE HCV NS4-NS5 peptide 2918.
KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PD 16-DEC-1992.
PF 11-JUN-1992; 109812.
PR 11-JUN-1991; JP-139268.
PR 12-JUL-1991; JP-172794.
PR 07-OCT-1991; JP-287008.
PR 16-DEC-1991; JP-332329.
PR 20-APR-1992; JP-099957.
PA (MITU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI; 92-417213/51.
DR N-PSDB; Q32510.
PT New hepatitis C virus gene and its encoded protein - used for
PT diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure; Page 207-08; 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequencess were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in Q32565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 285 AA;

R29879 Length: 285 February 11, 2000 15:48 Type: P Check: 8945 ..

1 YRDVLKDMKA KASTVKAKLL SVEEACKLTP PHSARSKFDY GAKDVQSLSS

51 KAVNHIHSVW KDLPEDTETP IDTTIMAKNE VFCVQPEKGG RKPARLIVFP

101 DLGVRVCEKM ALYDVVSTLP QAVMGSSYRF QCSPGQRVEF LVNAWKSKKS

151 PMGFAYDTRC FDSTVTENDI RTEESIYQCC DLDPEARQAI RSLTERLYIG

201 GPLTNSKGQN CGYRRCRASG VLTTSCGNTL TCYLKASAAC RAAKLQDCTM

251 LVCGDDLVVI CESAGTQEDA ANLRVFTEAM TRNSA

!!AA_SEQUENCE 1.0
ID R29895 standard; Protein; 171 AA.
AC R29895;
DT 26-APR-1993 (first entry)
DE HCV NS4-NS5 peptide N29-1.
KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PD 16-DEC-1992.
PF 11-JUN-1992; 109812.
PR 11-JUN-1991; JP-139268.
PR 12-JUL-1991; JP-172794.
PR 07-OCT-1991; JP-287008.
PR 16-DEC-1991; JP-332329.
PR 20-APR-1992; JP-099957.
PA (MITU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI; 92-417213/51.
DR N-PSDB; Q32526.
PT New hepatitis C virus gene and its encoded protein - used for
PT diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure; Page 242-43; 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequencess were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in Q32565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 171 AA;

R29895 Length: 171 February 11, 2000 15:48 Type: P Check: 8862 ..

1 YRDVLKEMKA KASTVKAKLL SVEEACKLTP PHSARSKFGY GAKDVRSLSS

51 KAVNHIRSVW KDLLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARLIVFP

101 DLGVRVCEKM ALYDVVSTLP QAVMGSSYGF QYSPGQRVEF LVNAWKSKKS

151 PMGFAYDTRC FDSTVTENDI R

!!AA_SEQUENCE 1.0
ID R29896 standard; Protein; 171 AA.
AC R29896;
DT 26-APR-1993 (first entry)
DE HCV NS4-NS5 peptide N29-2.

KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PD 16-DEC-1992.
PF 11-JUN-1992; 109812.
PR 11-JUN-1991; JP-139268.
PR 12-JUL-1991; JP-172794.
PR 07-OCT-1991; JP-287008.
PR 16-DEC-1991; JP-332329.
PR 20-APR-1992; JP-099957.
PA (MITU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI; 92-417213/51.
DR N-PSDB; Q32527.
PT New hepatitis C virus gene and its encoded protein - used for
PT diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure; Page 243-44; 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequencess were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in Q32565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 171 AA;

R29896 Length: 171 February 11, 2000 15:48 Type: P Check: 8862 ..

1 YRDVLKEMKA KASTVKAKLL SVEEACKLTP PHSARSKFGY GAKDVRSLSS

51 KAVNHIRSVW KDLLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARLIVFP

101 DLGVRVCEKM ALYDVVSTLP QAVMGSSYGF QYSPGQRVEF LVNAWKSKKS

151 PMGFAYDTRC FDSTVTENDI R

!!AA_SEQUENCE 1.0
ID R29897 standard; Protein; 167 AA.
AC R29897;
DT 26-APR-1993 (first entry)
DE HCV NS4-NS5 peptide N29-3.
KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele; core; region; upstream;
KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;
KW determinant; antiserum.
OS Hepatitis C virus
PN EP-518313-A.
PD 16-DEC-1992.
PF 11-JUN-1992; 109812.
PR 11-JUN-1991; JP-139268.
PR 12-JUL-1991; JP-172794.
PR 07-OCT-1991; JP-287008.
PR 16-DEC-1991; JP-332329.
PR 20-APR-1992; JP-099957.
PA (MITU) MITSUBISHI KASEI CORP.
PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;
PI Teranishi Y;
DR WPI; 92-417213/51.
DR N-PSDB; Q32528.
PT New hepatitis C virus gene and its encoded protein - used for
PT diagnosing and vaccinating against hepatitis C virus infections
PS Disclosure; Page 244-45; 305pp; English.
CC The sequences given in R29871-906 and R29533 are encoded by various
CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of
CC the invention. These NS4-NS5 RNA sequencess were isolated from the
CC serum of a patient suffering from hepatitis C (HC). The isolated RNA
CC sequences were converted into cDNA using transcriptase in the presence
CC of one of the primer sequences given in Q32565-77. The sequences were
CC then amplified using primer pairs. The cDNA sequences isolated
CC represent different alleles of the same region of the HCV gene.
CC Sequence analysis shows that these clones represent the core region
CC and some upstream sequences of HCV. These polypeptides are thought to
CC contain a highly hydrophilic region which can adopt a "turn structure"
CC which is not an alpha helix or a beta sheet. These polypeptides are
CC thought to act as antigen determinants and are highly reactive with
CC antiserum raised against HCV-associated antigens. See also Q32436.
SQ Sequence 167 AA;

R29897 Length: 167 February 11, 2000 15:48 Type: P Check: 2886 ..

1 YRDVLKEMKA KASTVKAKLL SVEEACKLTP PHSARSKFGY GAKDVRSLSS

51 KAVNHIRSVW KDLLEDTETP IDTTIMAKNE VFCVQPEKGG RKPARLIVFP

101 DLGVRVCEKM ALYDVVSTLP QAVMGSSYGF QYSPGQRVEF LVNAWKSKKS

151 PMGFSYDTRC FDSTVTE

!!AA_SEQUENCE 1.0
ID R29909 standard; Protein; 277 AA.
AC R29909;
DT 05-MAY-1993 (first entry)
DE Prod. of the luffin-f gene.
KW Loofah; amplification; missile treatment; virus; agriculture.
OS Luffa cylindrica.
PN J04320687-A.
PD 11-NOV-1992.
PF 18-APR-1991; 086857.
PR 18-APR-1991; JP-086857.
PA (NISB) JAPAN TOBACCO INC.
DR WPI; 92-426681/52.
DR N-PSDB; Q32687.
PT New gene coding luffin - for missile treatment and prevention of
PT viral diseases in agriculture
PS Disclosure; Page 10; 13pp; Japanese.
CC The sequence is that of the prod. of the luffin-f gene of Luffa
CC cylindrica. The gene sequence was obtd. by PCR of luffa cDNA
CC using a PCR primer designed based on the luffin protein sequence
CC from amino acid residues 54-60. The luffin gene can be used in
CC missile treatment and for prevention of viral diseases in agriculture.
CC Luffin may be produced stably, by transforming a microbe such as E. coli
CC with the luffin gene.
CC See also R22910.
SQ Sequence 277 AA;

R29909 Length: 277 February 11, 2000 15:48 Type: P Check: 429 ..

1 MKRFTVLILA IFVAASTVEA DVRFSLSGSS STSYSKFIGD LRKALPSNGT

51 VYNITLLLSS ASGASRYTLM TLSNYDGKAI TVAVDVTNVY IMGYLVNSTS

101 YFFNESDAKL ASQYVFKGST IVTLPYSGNY EKLQTAAGKI REKIPLGFPA

151 LDSAITTLFH YDSTAAAAAF LVIIQTTAEA SRFKYIEGQI IERISKNQVP

201 SLATISLENE WSALSKQIQL AQTNNGTFKT PVVITDDKGQ RVEITNVTSK

251 VVTKNIQLLL NYKQNVAAFD EDVSAKH

!!AA_SEQUENCE 1.0
ID R29910 standard; Protein; 278 AA.
AC R29910;
DT 05-MAY-1993 (first entry)
DE Prod. of the luffin-g gene.
KW Loofah; amplification; missile treatment; virus; agriculture.

OS Luffa cylindrica.
PN J04320687-A.
PD 11-NOV-1992.
PF 18-APR-1991; 086857.
PR 18-APR-1991; JP-086857.
PA (NISB) JAPAN TOBACCO INC.
DR WPI; 92-426681/52.
DR N-PSDB; Q32688.
PT New gene coding luffin - for missile treatment and prevention of
PT viral diseases in agriculture
PS Disclosure; Page 12; 13pp; Japanese.
CC The sequence is that of the prod. of the luffin-g gene of Luffa
CC cylindrica. The gene sequence was obtd. by PCR of luffa cDNA
CC using a PCR primer designed based on the luffin protein sequence
CC from amino acid residues 54-60. The luffin gene can be used in
CC missile treatment and for prevention of viral diseases in agriculture.
CC Luffin may be produced stably, by transforming a microbe such as E. coli
CC with the luffin gene.
CC See also R22909.
SQ Sequence 278 AA;

R29910 Length: 278 February 11, 2000 15:48 Type: P Check: 9157 ..

1 MNRFTFLSLL IILIAFFTVE GANVSFSLSG ADSKSYSKFI TALRKALPSK

51 EKVSNIPLLL PSASGASRYI LMQLSNYDAK AITMAIDVTN VYIMGYLVNS

101 TSYFFNESDA KLASQYVFKG STIVTLPYSG NYERLQNAAG KVREKIPLGF

151 RAFDSAITSL FHYDSTAAAG AFLVIIQTTA EASRFKYIEG QIIERIPKNE

201 VPSPAALSLE NEWSALSKQI QLAQTNNGAF RTPVVIIDNK GQRVEIKDVN

251 SKVVTNNIKL LLNKQNIAAF DDGIPTKH

!!AA_SEQUENCE 1.0
ID R31608 standard; Protein; 205 AA.
AC R31608;
DT 24-MAY-1993 (first entry)
DE Homologous to chicken nov gene exon 3-4-encoded protein.
KW avian nephroblastoma; avian myeloblastoma virus;
KW stringent hybridisation; ss.
PN WO9300430-A.
PD 07-JAN-1993.
PF 25-JUN-1992; F00589.
PR 25-JUN-1991; FR-007807.
PA (CNRS) CENT NAT RECH SCI.
PI Martinerie C, Perbal B;
DR WPI; 93-036377/04.
PT Nucleotide sequences hybridising to regions of chicken nov gene -
PT useful as probes for detecting complementary sequences to
PT evaluate development and/or differentiation of tumours
PS Claim 21; Page 39; 67pp; French.
CC The chicken nov gene is stimulated in avian nephroblastoma induced
CC by avian myeloblastoma virus but not in normal adult kidney. A
CC 1975bp cDNA sequence (Q36031) was isolated from a gene bank prepared
CC from chicken embryonic fibroblasts screened with a tumour-derived
CC probe. Fragment XXII (Q36044) is part of the 3rd and 4th exons of
CC the nov gene; nucleotide sequences which hybridise to Fragment XXII
CC under stringent conditions (i.e. 50% formamide, 5 x SCC) are claimed.
CC The claimed sequences preferably encode a protein with the sequence
CC XXIII (R31608) or an amino acid sequence 60% homologous to it.
SQ Sequence 205 AA;

R31608 Length: 205 February 11, 2000 15:48 Type: P Check: 7012 ..

1 QIPTRIPDAL DVRVPQCLTS ASPTPLFPSS SPAKDGAPCI FGGTVYRSGE

51 SFQSSCKYQC TCLDGAVGCM PLCSMDVRLP SPDCPFPRRV KLPGKCCEEW

101 VCDEPKDQTV LGPASRVSRV FLXVRVVILS QGGSPNCADR TGEIPYPGVD

151 HGVCVLCSRS LPTGRHVWPR PNYDXSQLPG PDTEWSACSK TCGMGYSTRV

201 TNDNA

!!AA_SEQUENCE 1.0
ID R30618 standard; Protein; 106 AA.
AC R30618;
DT 19-MAY-1993 (first entry)
DE Polypeptide coded by Korean HCV cDNA fragment NS2-LBC3.
KW KHCV; diagnosis; vaccine; NS2 region.
OS Korean hepatitis C virus.
PN EP-521318-A.
PD 07-JAN-1993.
PF 10-JUN-1992; 109753.
PR 10-JUN-1991; KR-009510.
PR 06-AUG-1991; KR-013601.
PA (LUCK-) LUCKY LTD.
PI Cho JM, Choi DY, Kim CH, Kim ST, Lee YB, Lim KJ, Park YW,
PI So HS, Yang JY;.
DR WPI; 93-001883/01.
DR N-PSDB; Q33284.
PT DNA and polypeptide(s) from a new type of hepatitis C virus (KHCV)
PT - for diagnosing and vaccinating against KHCV infections
PS Disclosure; Fig 8; 119pp; English.
CC The polypeptide is that encoded by cDNA fragment NS2-LBC3 obtd. from
CC Korean hepatitis C virus (KHCV) RNA extracted from the sera of
CC hepatitis C patients. It or its fragments may be used in a specific
CC and accurate method for detecting KCHV antibodies in the serum of
CC hepatitis C patients. Antibodies directed against the polypeptide
CC are useful for the purification of KCHV antigens and for the
CC development of an improved diagnostic to detect KHCV antigens in a
CC sample. The polypeptide may also be used in a vaccine for treatment
CC and prevention of KHCV infection at a dosage of 5-200 ug/peptide.
SQ Sequence 106 AA;

R30618 Length: 106 February 11, 2000 15:48 Type: P Check: 7998 ..

1 LLTLSPYYKV FLARLIWWLQ YFITRAEAHL QVWIPPLNVR GGRDAIILLA

51 CAVHPEPIFD ITKYLLAIFG PLMVLQAGIT RVPYFWRAQG LIRACMLARK

101 VAGGHY

!!AA_SEQUENCE 1.0
ID R32192 standard; Protein; 246 AA.
AC R32192;
DT 30-MAY-1993 (first entry)
DE Sequence encoded by the acetoacetyl-CoA reductase (phbB) gene of the
DE polyhydroxybutyrate (PHB) operon
KW Operon; polyhydroxyalkanoate; acetoacetyl-CoA reductase.
OS Alcaligenes eutrophus.
PN WO9302187-A.
PD 04-FEB-1993.
PF 13-JUL-1992; U05786.
PR 19-JUL-1991; US-732243.
PA (UYMA-) UNIV MADISON JAMES.
PA (UNMS) UNIV MICHIGAN STATE.
PI Dennis DE, Poirier Y, Somerville CR;
DR WPI; 93-058785/07.
DR N-PSDB; Q36660.
PT Transgenic plants producing poly:hydroxy-alkanoate polymer(s) -
PT obtd. by transformation with DNA encoding 3-ketothiolase,
PT acetoacetyl-CoA reductase and PHA synthase
PS Disclosure; Fig 2; 70pp; English.
CC The nucleotide sequence of the PHB operon was obtained from Janes, B.
CC Hollar, J. and Dennis, D. in Dawes, E.A. (ed.) Novel Biodegradable
CC Polymers, Kluwer Academic Publishers, 175-190 (1990). It contains
CC the genes from PHB synthase, 3-ketothiolase and acetoacetyl-CoA
CC reductase. The inventors claim a transgenic plant material contg.
CC foreign DNA encoding a peptide which exhibits 3-ketothilase activity,
CC pref. where the DNA is an open reading from between nucleotides
CC 2696-3877 (phb A gene), 842-2611 (phb C gene) or 3952-4692 (phb B
CC gene) of the Alcaligenes eutrophus PHB operon.
SQ Sequence 246 AA;

R32192 Length: 246 February 11, 2000 15:48 Type: P Check: 2829 ..

```
       1  MTQRIAYVTG GMGGIGTAIC QRLAKDGFRV VAGCGPNSPR REKWLEQQKA

      51  LGFDFIASEG NVADWDSTKT AFDKVKSEVG EVDVLINNAG ITRDVVFRKM

     101  TRADWDAVID TNLTSLFNVT KQVIDGMADR GWGRIVNISS VNGQKGQFGQ

     151  TNYSTAKAGL HGFTMALAQE VATKGVTVNT VSPGYIATDM VKAIRQDVLD

     201  KIVATIPVKR LGLPEEIASI CAWLSSEESG FSTGADFSLN GGLHMG
```

```
!!AA_SEQUENCE 1.0
ID   R37294 standard; protein; 248 AA.
AC   R37294;
DT   13-SEP-1993  (first entry)
DE   Plant type I RIP Luffin.
KW   Type I ribosome-inactivating protein; ricin; gelonin; momordin;
KW   immunoconjugate; autoimmune disease; cell killing; toxin; loofah.
OS   Luffa.
PN   WO9309130-A.
PD   13-MAY-1993.
PF   04-NOV-1992; U09487.
PR   04-NOV-1991; US-787567.
PR   19-JUN-1992; US-901707.
PA   (XOMA ) XOMA CORP.
PI   Berhard SL, Better MD, Carroll SF, Lane JA, Lei SP;
DR   WPI; 93-167617/20.
PT   Analogues of type I ribosome inactivating protein - useful as
PT   cytotoxic agents, immuno toxins for treating auto immune diseases,
PT   cancer, graft versus host disease and selective cell killing in-vivo
PS   Disclosure; Page 95-96; 163pp; English.
CC   The invention covers analogues of Type I RIPs. Luffin is a
CC   Type I RIP and the analogues of the invention have a cysteine
CC   available for intermolecular disulphide bonding at an amino acid
CC   position corresp. to a position not naturally available for bonding;
CC   the cys residue is located in the C-terminal region of the analogue
CC   between a position corresp. to amino acid 251 and the C-terminus of
CC   ricin A chain. The analogues are pref. joined via a disulphide
CC   linkage to a molecule which specifically binds to a target cell, e.g.
CC   an antibody fragment.
SQ   Sequence   248 AA;
```

R37294 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..

```
       1  DVRFSLSGSS STSYSKFIGD LRKALPSNGT VYNLTILLSS ASGASRYTLM

      51  TLSNYDGKAI TVAVDVSQLY IMGYLVNSTS YFFNESDAKL ASQYVFKGST

     101  IVTLPYSGNY EKLQTAAGKI REKIPLGFPA LDSALTTIFH YDSTAAAAAF

     151  LVILQTTAEA SRFKYIEGQI IERISKNQVP SLATISLENS LWSALSKQIQ

     201  LAQTNNGTFK TPVVITDDKG QRVEITNVTS KVVTKNIQLL LNYKQNVA
```

```
!!AA_SEQUENCE 1.0
ID   R39259 standard; Protein; 391 AA.
AC   R39259;
DT   18-NOV-1993  (first entry)
DE   Human somatostatin receptor-1.
KW   Human; somatostatin; receptor; SSTR-1; SSTR-2; SSTR-3; tumour;
KW   pancreas; islet; promoter; transformation; host cell.
OS   Homo sapiens.
PN   WO9313130-A.
PD   08-JUL-1993.
PF   30-DEC-1992; U11327.
PR   31-DEC-1991; US-816283.
PA   (ARCH-) ARCH DEV CORP.
PI   Bell GI,  Seino S,  Yamada Y;
DR   WPI; 93-227272/28.
DR   N-PSDB; Q45653.
PT   Somatostatin receptors useful for diagnosis of tumours - also
PT   useful for screening candidate somatostatin receptor agonists and
PT   antagonists
PS   Claim 3; Page 63-64; 94pp; English.
CC   The sequences given in R39259, R39261 and R39263 represent the human
CC   somatostatin receptors (SSTR)-1, SSTR-2 and SSTR-3.  The DNA encoding
CC   these proteins was isolated from total human pancreatic islet DNA.
CC   These DNA sequences may be placed under the control of a suitable
CC   promoter and used to transform a host cell.  The DNA sequences and
CC   these proteins may be used in screening assays for testing candidates
CC   including agonists and antagonists of SSTR polypeptides.  The assays
CC   may be used to discriminate candidate substances with desirable
CC   properties specific to SSTR polypeptides.  The isolated substances
CC   may be used in a wide range of aplications eg. diagnosis of various
CC   human tumours.  Fragments of these DNA sequences may be used as
CC   probes in the isolation of other SSTR-encoding clones.
SQ   Sequence   391 AA;
```

R39259 Length: 391 February 11, 2000 15:48 Type: P Check: 8652 ..

```
       1  MFPNGTASSP SSSPSPSPGS CGEGGGSRGP GAGAADGMEE PGRNASQNGT

      51  LSEGQGSAIL ISFIYSVVCL VGLCGNSMVI YVILRYAKMK TATNIYILNL

     101  AIADELLMLS VPFLVTSTLL RHWPFGALLC RLVLSVDAVN MFTSIYCLTV

     151  LSVDRYVAVV HPIKAARYRR PTVAKVVNLG VWVLSLLVIL PIVVFSRTAA

     201  NSDGTVACNM LMPEPAQRWL VGFVLYTFLM GFLLPVGAIC LCYVLIIAKM

     251  RMVALKAGWQ QRKRSERKIT LMVMMVVMVF VICWMPFYVV QLVNVFAEQD

     301  DATVSQLSVI LGYANSCANP ILYGFLSDNF KRSFQRILCL SWMDNAAEEP

     351  VDYYATALKS RAYSVEDFQP ENLESGGVFR NGTCTSRITT L
```

```
!!AA_SEQUENCE 1.0
ID   R39260 standard; Protein; 391 AA.
AC   R39260;
DT   18-NOV-1993  (first entry)
DE   Murine somatostatin receptor-1.
KW   Mouse; somatostatin; receptor; SSTR-1; SSTR-2; SSTR-3; tumour;
KW   pancreas; islet; promoter; transformation; host cell.
OS   Mus musculus.
PN   WO9313130-A.
PD   08-JUL-1993.
PF   30-DEC-1992; U11327.
PR   31-DEC-1991; US-816283.
PA   (ARCH-) ARCH DEV CORP.
PI   Bell GI,  Seino S,  Yamada Y;
DR   WPI; 93-227272/28.
DR   N-PSDB; Q45654.
PT   Somatostatin receptors useful for diagnosis of tumours - also
PT   useful for screening candidate somatostatin receptor agonists and
PT   antagonists
PS   Claim 3; Page 65-66; 94pp; English.
CC   The sequences given in R39260, R39262 and R39264 represent the murine
CC   somatostatin receptors (SSTR)-1, SSTR-2 and SSTR-3.  The DNA encoding
CC   these proteins was isolated from total murine pancreatic islet DNA.
CC   These DNA sequences may be placed under the control of a suitable
CC   promoter and used to transform a host cell.  The DNA sequences and
CC   these proteins may be used in screening assays for testing candidates
CC   including agonists and antagonists of SSTR polypeptides.  The assays
CC   may be used to discriminate candidate substances with desirable
CC   properties specific to SSTR polypeptides.  The isolated substances
CC   may be used in a wide range of aplications eg. diagnosis of various
CC   human tumours.  Fragments of these DNA sequences may be used as
CC   probes in the isolation of other SSTR-encoding clones.
SQ   Sequence   391 AA;
```

R39260 Length: 391 February 11, 2000 15:48 Type: P Check: 8110 ..

```
       1  MFPNGTASSP SSSPSPSPGS CGEGACSRGP GSGAADGMEE PGRNASQNGT
```

```
     51  LSEGQGSAIL ISFIYSVVCL VGLCGNSMVI YVILRYAKMK TATNIYILNL

    101  AIADELLMLS VPFLVTSTLL RHWPFGALLC RLVLSVDAVN MFTSIYCLTV

    151  LSVDRYVAVV HPIKAARYRR PTVAKVVNLG VWVLSLLVIL PIVVFSRTAA

    201  NSDGTVACNM LMPEPAQRWL VGFVLYTFLM GFLLPVGAIC LCYVLIIAKM

    251  RMVALKAGWQ QRKRSERKIT LMVMMVVMVF VICWMPFYVV QLVNVFAEQD

    301  DATVSQLSVI LGYANSCANP ILYGFLSDNF KRSFQRILCL SWMDNAAEEP

    351  VDYYATALKS RAYSVEDFQP ENLESGGVFR NGTCASRIST L

!!AA_SEQUENCE 1.0
ID   R42317 standard; Protein; 345 AA.
AC   R42317;
DT   22-MAY-1994  (first entry)
DE   EBV VCA-p40 from ORF BdRF1.
KW   Viral capsid antigen; Epstein Barr Virus; viral genome; antibodies;
KW   diagnosis.
OS   Epstein-barr virus.
PN   AU9335152-A.
PD   16-SEP-1993.
PF   12-MAR-1993; 035152.
PR   13-MAR-1992; EP-200721.
PA   (ALKU ) AKZO NV.
PI   Middeldorp JM;
DR   WPI; 93-345368/44.
DR   N-PSDB; Q51019.
PT   Peptide(s) and nucleic acid sequences related to EBV - for
PT   detecting EBV in samples
PS   Claim 3; Page 40; 62pp; English.
CC   The sequence is that of Epstein Barr Virus viral capsid antigen
CC   VCA-p40 encoded by the EBV open reading frame BdRF1. This antigen is
CC   immunochemically reactive with antibodies to EBV and can be used for
CC   the diagnosis of EBV in a sample.
CC   See also R42316-35.
SQ   Sequence   345 AA;

 R42317  Length: 345  February 11, 2000 15:48  Type: P  Check: 3489  ..

       1  MLSGNAGEGA TACGGSAAAG QDLISVPRNT FMTLLQTNLD NKPPRQTPLP

      51  YAAPLPPFSH QAIATAPSYG PGAGAVAPAG GYFTSPGGYY AGPAGGDPGA

     101  FLAMDAHTYH PHPHPPPAYF GLPGLFGPPP PVPPYYGSHL RADYVPAPSR

     151  SNKRKRDPEE DEEGGGLFPG EDATLYRKDI AGLSKSVNEL QHTLQALRRE

     201  TLSYGHTGVG YCPQQGPCYT HSGPYGFQPH QSYEVPRYVP HPPPPPTSHQ

     251  AAQAQPPPPG TQAPEAHCVA ESTIPEAGAA RNSGPREDTN PQQPTTEGHH

     301  RGKKLVQASA SGVAQSKEPT TPKAKSVSAH LKSIFCEELL NKRVA

!!AA_SEQUENCE 1.0
ID   R52667 standard; Protein; 303 AA.
AC   R52667;
DT   09-AUG-1994  (first entry)
DE   Equine herpesvirus US2 gene.
KW   Equine herpesvirus; US2; vaccine; antigen; protection; prophylaxis;
KW   prevention.
OS   Equine herpesvirus (Strain: Dutta; Individual isolate: S-1EHV-000).
PN   WO9403628-A.
PD   17-FEB-1994.
PF   06-AUG-1993; U07424.
PR   07-AUG-1992; US-926784.
PA   (SYTR ) SYNTRO CORP.
PI   Cochran MD;
DR   WPI; 94-065715/08.
DR   N-PSDB; Q56616.
PT   New recombinant equine herpes viruses - used to prepare vaccines
PT   to protect horses from infectious equine herpes virus
PS   Example 1; Page 83-84; 153pp; English.
CC   Recombinant equine herpesvirus' (EHV) have a foreign DNA sequence
CC   inserted into their genomes.  The inserted sequence is a piece of
CC   foreign, double stranded DNA which encodes an RNA molecule which
CC   does not naturally occur in the animal into which the EHV is
CC   introduced.  The foreign DNA sequence is introduced into the EHV via
CC   a homology vector.  The recombinant EHV's are used to prepare
CC   vaccines to protect horses from infectious EHV.  This is the US2
CC   gene product from a recombinant EHV.  Recombinant attenuated EHV's
CC   can be produced by removing the US2 gene
SQ   Sequence   303 AA;

 R52667  Length: 303  February 11, 2000 15:48  Type: P  Check: 5723  ..

       1  MGVVLITVVT VVDRHKALPN SSIDVDGHLW EFLSRQCFVL ASEPLGIPIV

      51  VRSADLYRFS SSLLTLPKAC RPIVRTRGAT AIALDRNGVV YHEDRMGVSI

     101  EWLSVLSGYN HLNSSLIINQ PYHLWVLGAA DLCKPVFDLI PGPKRMVYAE

     151  IADEFHKSWQ PPFVCGKLFE TIPWTTVEHN HPLKLRAAGG EDTVVGECGF

     201  SKHSSNSLVR PPTVKRVIYA VVDPARLREI PAPGRPLPRR RPSEGGMRAP

     251  RRRSRAPAAA RSTAAAATPP RPGDPRAPAA RRAGDVTWME RLLWGVFGRT

     301  STR

!!AA_SEQUENCE 1.0
ID   R49296 standard; Protein; 25 AA.
AC   R49296;
DT   16-SEP-1994  (first entry)
DE   Invarient chain (Ii) position 97-121.
KW   Naturally-occuring; immunomodulatory protein; human; therapy; class I;
KW   major histocompatibility complex; class II; allotype; type I diabetes;
KW   autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW   mutiple sclerosis; transplant rejection; vaccine; MHC.
OS   Homo sapiens.
PN   WO9404171-A.
PD   03-MAR-1994.
PF   11-AUG-1993; U07545.
PR   11-AUG-1992; US-925460.
PR   15-JUN-1993; US-925460.
PA   (HARD ) HARVARD COLLEGE.
PI   Chicz RM,  Hedley ML,  Stern LJ,  Strominger JL,  Urban RG;
PI   Vignali DA;
DR   WPI; 94-082825/10.
PT   Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT   treatment of auto:immune diseases, transplant rejection and for
PT   vaccination
PS   Disclosure; Page 36; 139pp; English.
CC   The sequences given in R49291-505 and R46981-7038 represent peptide
CC   fragments of naturally-occuring immunomodulatory proteins.  These
CC   fragments are between 10-30 residues in length and bind to a human
CC   major histocompatibility complex (MHC) class II allotype.  These
CC   peptides may be used for therapy of autoimmune diseases, such as
CC   type I diabetes, rheumatoid arthritis and mutiple sclerosis, and to
CC   reduce transplant rejection.  They may also be used for vaccination
CC   providing an exclusively T-cell-mediated response, which can be
CC   class I or class-II based, or both, depending on the length and
CC   character of the immunogenic peptides.
SQ   Sequence   25 AA;

 R49296  Length: 25  February 11, 2000 15:48  Type: P  Check: 4901  ..

       1  LPKPPKPVSK MRMATPLLMQ ALPMG

!!AA_SEQUENCE 1.0
ID   R49297 standard; Protein; 24 AA.
AC   R49297;
DT   16-SEP-1994  (first entry)
DE   Invarient chain (Ii) position 97-120.
```

KW Naturally-occuring; immunomodulatory protein; human; therapy; class I;
KW major histocompatibility complex; class II; allotype; type I diabetes;
KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW mutiple sclerosis; transplant rejection; vaccine; MHC.
OS Homo sapiens.
PN WO9404171-A.
PD 03-MAR-1994.
PF 11-AUG-1993; U07545.
PR 11-AUG-1992; US-925460.
PR 15-JUN-1993; US-925460.
PA (HARD) HARVARD COLLEGE.
PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
PI Vignali DA;
DR WPI; 94-082825/10.
PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT treatment of auto:immune diseases, transplant rejection and for
PT vaccination
PS Disclosure; Page 36; 139pp; English.
CC The sequences given in R49291-505 and R46981-7038 represent peptide
CC fragments of naturally-occuring immunomodulatory proteins. These
CC fragments are between 10-30 residues in length and bind to a human
CC major histocompatibility complex (MHC) class II allotype. These
CC peptides may be used for therapy of autoimmune diseases, such as
CC type I diabetes, rheumatoid arthritis and mutiple sclerosis, and to
CC reduce transplant rejection. They may also be used for vaccination
CC providing an exclusively T-cell-mediated response, which can be
CC class I or class-II based, or both, depending on the length and
CC character of the immunogenic peptides.
SQ Sequence 24 AA;

R49297 Length: 24 February 11, 2000 15:48 Type: P Check: 3126 ..

1 LPKPPKPVSK MRMATPLLMQ ALPM

!!AA_SEQUENCE 1.0
ID R49299 standard; Protein; 23 AA.
AC R49299;
DT 16-SEP-1994 (first entry)
DE Invarient chain (Ii) position 97-119.
KW Naturally-occuring; immunomodulatory protein; human; therapy; class I;
KW major histocompatibility complex; class II; allotype; type I diabetes;
KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW mutiple sclerosis; transplant rejection; vaccine; MHC.
OS Homo sapiens.
PN WO9404171-A.
PD 03-MAR-1994.
PF 11-AUG-1993; U07545.
PR 11-AUG-1992; US-925460.
PR 15-JUN-1993; US-925460.
PA (HARD) HARVARD COLLEGE.
PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
PI Vignali DA;
DR WPI; 94-082825/10.
PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT treatment of auto:immune diseases, transplant rejection and for
PT vaccination
PS Disclosure; Page 36; 139pp; English.
CC The sequences given in R49291-505 and R46981-7038 represent peptide
CC fragments of naturally-occuring immunomodulatory proteins. These
CC fragments are between 10-30 residues in length and bind to a human
CC major histocompatibility complex (MHC) class II allotype. These
CC peptides may be used for therapy of autoimmune diseases, such as
CC type I diabetes, rheumatoid arthritis and mutiple sclerosis, and to
CC reduce transplant rejection. They may also be used for vaccination
CC providing an exclusively T-cell-mediated response, which can be
CC class I or class-II based, or both, depending on the length and
CC character of the immunogenic peptides.
SQ Sequence 23 AA;

R49299 Length: 23 February 11, 2000 15:48 Type: P Check: 1278 ..

1 LPKPPKPVSK MRMATPLLMQ ALP

!!AA_SEQUENCE 1.0
ID R49438 standard; Protein; 49 AA.
AC R49438;
DT 16-SEP-1994 (first entry)
DE Minigene HLA-DRalpha chain/Ii 24 residue peptide fusion.
KW Naturally-occuring; immunomodulatory protein; human; therapy; class I;
KW major histocompatibility complex; class II; allotype; type I diabetes;
KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW mutiple sclerosis; transplant rejection; vaccine; MHC.
OS Homo sapiens.
PN WO9404171-A.
PD 03-MAR-1994.
PF 11-AUG-1993; U07545.
PR 11-AUG-1992; US-925460.
PR 15-JUN-1993; US-925460.
PA (HARD) HARVARD COLLEGE.
PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
PI Vignali DA;
DR WPI; 94-082825/10.
PT Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT treatment of auto:immune diseases, transplant rejection and for
PT vaccination
PS Disclosure; Page 93; 139pp; English.
CC The sequences given in R49291-505 and R46981-7038 represent peptide
CC fragments of naturally-occuring immunomodulatory proteins. These
CC fragments are between 10-30 residues in length and bind to a human
CC major histocompatibility complex (MHC) class II allotype. These
CC peptides may be used for therapy of autoimmune diseases, such as
CC type I diabetes, rheumatoid arthritis and mutiple sclerosis, and to
CC reduce transplant rejection. They may also be used for vaccination
CC providing an exclusively T-cell-mediated response, which can be
CC class I or class-II based, or both, depending on the length and
CC character of the immunogenic peptides.
SQ Sequence 49 AA;

R49438 Length: 49 February 11, 2000 15:48 Type: P Check: 4300 ..

1 MAISGVPVLG FFIIAVLMSA QESWALPKPP KPVSKMRMAT PLLMQALPM

!!AA_SEQUENCE 1.0
ID R49589 standard; Protein; 49 AA.
AC R49589;
DT 15-SEP-1994 (first entry)
DE Sequence of HLA-DR alpha chain leader peptide linked to the
DE amino terminus of a blocking peptide fragment of human invariant
DE chain Ii.
KW Minigene; HLA-DR alpha leader peptide; chain Ii.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1. .25
FT /label= HLA-DR alpha leader peptide
FT peptide 26. .49
FT /label= chain Ii
PN WO9404557-A.
PD 03-MAR-1994.
PF 11-AUG-1992; U06692.
PR 11-AUG-1992; WO-U06692.
PA (HARD) HARVARD COLLEGE.
PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;
PI Vignali DAA;
DR WPI; 94-083102/10.
DR N-PSDB; Q44334.
PT New peptide binding to MHC class II allotype - useful for
PT treating auto-immune diseases, transplant rejection and for
PT immunisation
PS Example; Figure 3B; 60pp; English.
CC The inventors claim a nucleic acid encoding a polypeptide
CC comprising an AA sequence identical to that of a segment of a
CC naturally-occurring human protein which binds to a human
CC major histocompatibility complex (MHC) class II allotype and
CC a second sequence substantially identical to human Ii. In order
CC to construct such minigenes, overlapping synthetic oligos
CC were used to generate the leader peptide/blocking peptide

CC minigenes (see Q44333/R49588, Q44334/R49589). The resulting
CC constructs were cloned into pGEM-2.
SQ Sequence 49 AA;

R49589 Length: 49 February 11, 2000 15:48 Type: P Check: 4300 ..

1 MAISGVPVLG FFIIAVLMSA QESWALPKPP KPVSKMRMAT PLLMQALPM

!!AA_SEQUENCE 1.0
ID R50081 standard; Protein; 352 AA.
AC R50081;
DT 24-NOV-1994 (first entry)
DE NANBH virus antigenic fragment #13.
KW Antigen; structural; non-structural; non A non B hepatitis virus;
KW NANBHV; NANBH; patient; plasma; diagnosis; detection; carrier; ss.
OS Non A Non B hepatitis virus.
PN J06070778-A.
PD 15-MAR-1994.
PF 01-JUN-1993; 156087.
PR 10-JUL-1992; JP-207391.
PA (KOKU-) KOKUSAI SHIYAKU KK.
PA (SANW) SANWA KAGAKU KENKYUSHO CO.
PA (TOFU) TONEN CORP.
PA (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
DR WPI; 94-128677/16.
DR N-PSDB; Q58826.
PT Nucleic acid fragment coding non-A non-B hepatitis virus antigen
PT - useful in diagnosis of NANB patient and detection of virus
PT carrier
PS Claim 26; Page 31-32; 37pp; Japanese.
CC The sequences given in R50068 and R50070-82 represent antigens of
CC structural and non-structural regions of non A non B hepatitis virus
CC (NANBHV). The cDNA encoding these sequences wase derived from the
CC plasma of a NANBH patient by recombinant DNA techniques. These
CC fragments are useful for the diagnosis of NANBH patients and the
CC detection of NANBHV carriers.
SQ Sequence 352 AA;

R50081 Length: 352 February 11, 2000 15:48 Type: P Check: 3747 ..

1 LLRYHNKVYC TTSKSASLRA KKVTFDRMQV LDAHYDSVLK DIKLAASKVS

51 ARLLTLEEAC RLTPPHSARS KYGFGAKEVR SLSGRAVNHI KSVWKDLLED

101 PQTPIPTTIM AKNEVFCVDP TKGGKKPARL IVYPDLGVRV CEKMALYDVT

151 QKLPQAVMGA SYGFQYSPAQ RVEFLLKAWA DKKDPMGFSY DTRCFDSTVT

201 ERDIRTEESI YQACSLPEEA RTAIRSLTER LYVGGPMFNS KGQACGYRRC

251 RASGVLTTSM GNTITCYVKA LAACKAAGIV APTMLVCGDD LVVISESQGT

301 EEDERNLRAF TEAMTRYSAP PGDPPRPEYD LELITSCSSN VSVAISPQGR

351 RR

!!AA_SEQUENCE 1.0
ID R50082 standard; Protein; 215 AA.
AC R50082;
DT 24-NOV-1994 (first entry)
DE NANBH virus antigenic fragment #14.
KW Antigen; structural; non-structural; non A non B hepatitis virus;
KW NANBHV; NANBH; patient; plasma; diagnosis; detection; carrier; ss.
OS Non A Non B hepatitis virus.
PN J06070778-A.
PD 15-MAR-1994.
PF 01-JUN-1993; 156087.
PR 10-JUL-1992; JP-207391.
PA (KOKU-) KOKUSAI SHIYAKU KK.
PA (SANW) SANWA KAGAKU KENKYUSHO CO.
PA (TOFU) TONEN CORP.
PA (TOKR-) ZH TOKYOTO RINSHO IGAKU SOGO KENKYUSHO.
DR WPI; 94-128677/16.
DR N-PSDB; Q58827.
PT Nucleic acid fragment coding non-A non-B hepatitis virus antigen
PT - useful in diagnosis of NANB patient and detection of virus
PT carrier
PS Claim 28; Page 32-33; 37pp; Japanese.
CC The sequences given in R50068 and R50070-82 represent antigens of
CC structural and non-structural regions of non A non B hepatitis virus
CC (NANBHV). The cDNA encoding these sequences wase derived from the
CC plasma of a NANBH patient by recombinant DNA techniques. These
CC fragments are useful for the diagnosis of NANBH patients and the
CC detection of NANBHV carriers.
SQ Sequence 215 AA;

R50082 Length: 215 February 11, 2000 15:48 Type: P Check: 108 ..

1 RYSAPPGDPP RPEYDLELIT SCSSNVSVAI SPQGRRRYYL SRDPTTPIAR

51 AAWETVRHSP VNSWLGNIIQ YAPTIWVRMV LMTHFFPVLI AQDTLDQNLN

101 FEMYGSVYSV SPLDLPAIIE RLHGLDAFSL HTYTHHELTR VASALRKLGA

151 PPLRAWKSRA RAVRASLISQ GGRAAVCGRY LFNWAVKTKL KLTPLPEARL

201 LDLSSWFTVG AGGGD

!!AA_SEQUENCE 1.0
ID R53251 standard; Protein; 56 AA.
AC R53251;
DT 09-DEC-1994 (first entry)
DE Consensus sequence of signal peptide of small subunit of RUBISCO.
KW Glycogen synthase; starch; gene expression; recombination;
KW transgenic plant; glucose-1-phosphate adenylyl transferase;
KW biosynthesis; metabolism; metabolic pathway.
OS Synthetic.
PN WO9409144-A.
PD 28-APR-1994.
PF 14-OCT-1992; G01881.
PR 14-OCT-1992; AU-026964.
PR 14-OCT-1992; WO-G01881.
PA (ZENE) ZENECA LTD.
PI Bird CR, Fentem PA, Keeling PL, Singletary G;
DR WPI; 94-151328/18.
PT New plants with altered starch synthesis ability - obtd. by
PT incorporating a donor gene specifying an enzyme involved in
PT starch or glycogen biosynthesis
PS Disclosure; Page 31; 91pp; English.
CC The glycogen synthase gene may be incorporated into a plant genome
CC to produce a plant with altered starch synthesising ability. Other
CC genes encoding enzymes involved in the starch or glycogen
CC biosynthetic pathways may also/additionally be used. The new plants
CC have an improved capacity to produce starch at elevated or lowered
CC temperatures and/or an ability to synthesise starch with an altered
CC fine structure. They can produce higher starch yields at certain
CC temperatures and/or produce starch with an improved quality. The
CC glycogen synthase gene construct used requires the presence of an
CC amyloplast transit peptide to ensure its correct localisation in the
CC amyloplast. This is the consensus sequence of the transit peptide of
CC the small subunit of RUBISCO from many genotypes.
SQ Sequence 56 AA;

R53251 Length: 56 February 11, 2000 15:48 Type: P Check: 2035 ..

1 MASSMLSSAA VATRTNPAQA SMVAPFTGLK SAAFPVSRKQ NLDITSIASN

51 GGRVQC

!!AA_SEQUENCE 1.0
ID R53646 standard; Protein; 380 AA.
AC R53646;
DT 18-JAN-1995 (first entry)
DE c-fos gene product.
KW c-jun; c-fos; cancer; tumour; metastasis; treatment; prophylaxis;
KW vaccine; prevention; lymphocytes; inhibition; antigen; antitumour.

OS Homo sapiens.
PN EP-599077-A.
PD 01-JUN-1994.
PF 28-OCT-1993; 117519.
PR 29-OCT-1992; US-968415.
PA (YEDA) YEDA RES & DEV CO LTD.
PI Eisenbach L, Feldmann M;
DR WPI; 94-169430/21.
DR N-PSDB; Q63815.
PT New antitumour vaccines partic. for cancer treatment - comprises
PT tumour cells having inserted c-fos and opt. c-jun genes, or
PT antigens expressed by these genes
PS Disclosure; Page 54-55; 68pp; English.
CC The c-fos gene can be transfected alone or co-transfected with the
CC c-jun gene (See Q63813) into tumour cells to provide a novel
CC anti-tumour cellular vaccine. The anti-tumour vaccine comprises
CC antigens expressed by the c-fos and c-jun genes. The vaccines
CC stimulate cytotoxic T-lymphocytes to produce an anti-tumourigenic and
CC anti-metastatic immune response. They can be used for contolling
CC tumourigenicity and metastatic properties of tumours. They are used
CC particularly for treating a patient suffering from cancer to prevent
CC and/or inhibit the development of metastasis.
SQ Sequence 380 AA;

R53646 Length: 380 February 11, 2000 15:48 Type: P Check: 4677 ..

1 MMFSGFNADY EASSSRCSSA SPAGDSLSYY HSPADSFSSM GSPVNAQDFC

51 TDLAVSSANF IPTVTAISTS PDLQWLVQPA LVSSVAPSQT RAPHPFGVPA

101 PSAGAYSRAG VVKTMTGGRA QSIGRRGKVE QLSPEEEEKR RIRRERNKMA

151 AAKCRNRRRE LTDTLQAETD QLEDEKSALQ TEIANLLKEK EKLEFILAAH

201 RPACKIPDDL GFPEEMSVAS LDLTGGLPEV ATPESEEAFT LPLLNDPEPK

251 PSVEPVKSIS SMELKTEPFD DFLFPASSRP SGSETARSVP DMDLSGSFYA

301 ADWEPLHSGS LGMGPMATEL EPLCTPVVTC TPSCTAYTSS FVFTYPEADS

351 FPSCAAAHRK GSSSNEPSSD SLSSPTLLAL

!!AA_SEQUENCE 1.0
ID R53748 standard; Protein; 355 AA.
AC R53748;
DT 06-FEB-1995 (first entry)
DE Seven transmembrane receptor (V28).
KW Primer; seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction.
OS Synthetic.
FH Key Location/Qualifiers
FT domain 26. .56
FT /label= Transmembrane domain.
FT domain 68. .92
FT /label= Transmembrane domain.
FT domain 107. .125
FT /label= Transmembrane domain.
FT domain 146. .167
FT /label= Transmembrane domain.
FT domain 197. .219
FT /label= Transmembrane domain.
FT domain 232. .253
FT /label= Transmembrane domain.
FT domain 273. .297
FT /label= Transmembrane domain.
PN WO9412635-A.
PD 09-JUN-1994.
PF 17-NOV-1993; U11153.
PR 17-NOV-1992; US-977452.
PA (ICOS-) ICOS CORP.
PI Godiska R, Gray PW, Schweickart VL;
DR WPI; 94-200264/24.
DR N-PSDB; Q66167.
PT DNA encoding seven trans:membrane receptors - used to develop
PT prods. for use as therapeutic or diagnostic agents for conditions
PT involving the receptors.
PS Example 6; Page 65-66; 100pp; English.
CC Two synthetic oligonucleotides (Q66165, Q66166) were annealed and
CC radiolabelled to produce a V28 specific probe. The probe was
CC denatured by boiling for two minutes and then hybridised to the
CC human placenta genomic library. Several strong hybridising signals
CC were identified and plaque purified and the DNA sequence (Q66167)
CC of a full length genomic clone encoding this seven transmembrane
CC receptor was deduced.
SQ Sequence 355 AA;

R53748 Length: 355 February 11, 2000 15:48 Type: P Check: 2863 ..

1 MDQFPESVTE NFEYDDLAEA CYIGDIVVFG TVFLSIFYSV IFAIGLVGNL

51 LVVFALTNSK KPKSVTDIYL LNLALSDLLF VATLPFWTHY LINEKGLHNA

101 MCKFTTAFFF IGFFGSIFFI TVISIDRYLA IVLAANSMNN RTVQHGVTIS

151 LGVWAAAILV AAPQFMFTKQ KENECLGDYP EVLQEIWPVL RNVETNFLGF

201 LLPLLIMSYC YFRIIQTLFS CKNHKKAKAI KLILLVVIVF FLFWTPYNVM

251 IFLETLKLYD FFPSCDMRKD LRLALSVTET VAFSHCCLNP LIYAFAGEKF

301 RRYLYHLYGK CLAVLCGRSV HVDFSSSESQ RSRHGSVLSS NFTYHTSDGD

351 ALLLL

!!AA_SEQUENCE 1.0
ID R58556 standard; protein; 175 AA.
AC R58556;
DT 28-MAR-1995 (first entry)
DE Heparin-binding secretory transforming factor 2.
KW heparin-binding growth factor family; fibroblast growth factor;
KW FGF; hst-2; heparin-binding secretory transforming factor 2;
KW drugs; reagents; growth promoting activity; increase platelet number;
KW angiogenesis promoting activity; treatment; burns; wounds; ulcers.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_difference 1
FT /note= "may be absent"
FT misc_difference 2. .30
FT /note= "N terminal fragment as given in the
FT specification"
PN EP-613946-A.
PD 07-SEP-1994.
PF 11-FEB-1994; 102093.
PR 12-FEB-1993; JP-024399.
PR 28-JUL-1993; JP-186507.
PA (NACA-) NAT CANCER CENT.
PA (TAKE) TAKEDA CHEM IND LTD.
PI Igarashi K, Ishimaru K, Terada M, Yoshimura K, Tereda M;
DR WPI; 94-272993/34.
PT New mutant hst-2 polypeptides - having cell growth promoting
PT activity used for treating eg. burns, wounds, post:operative
PT tissues or ulcers
PS Claim 15; Page 27; 39pp; English.
CC This sequence is a heparin-binding secretory transforming factor 2
CC (hst-2) polypeptide. It is preferred that the polypeptides of the
CC invention have regions deleted. The polypeptides have cell growth
CC promoting activity and angiogenesis promoting activity and can be
CC used for the treatment of burns, wounds, postoperative tissues or
CC ulcers. The therapeutic composition can be used as a platelet
CC increasing agent for use after radiotherapy and chemotherapy. (See
CC also R58553-5).
SQ Sequence 175 AA;

R58556 Length: 175 February 11, 2000 15:48 Type: P Check: 4528 ..

1 MLAGEIAGVN WESGYLVGIK RQRRLYCNVG IGFHLQVLPD GRISGTHEEN

```
     51  PYSLLEISTV ERGVVSLFGV RSALFVAMNS KGRLYATPSF QEECKFRETL

    101  LPNNYNAYES DLYQGTYIAL SKYGRVKRGS KVSPIMTVTH FLPRIVVPSP

    151  AGTRANNTLL DSRGWGTLLS RSRAG

!!AA_SEQUENCE 1.0
ID   R63592 standard; Peptide; 79 AA.
AC   R63592;
DT   09-MAY-1995  (first entry)
DE   MAP-kinase-phosphatase CL100 partial sequence.
KW   MAP-kinase inhibitor; MAPK; MAP-kinase-phosphatase; CL100;
KW   neurological disease; proliferative disorder; inflammation;
KW   cardiovascular disease.
OS   Homo sapiens.
PN   WO9423039-A.
PD   13-OCT-1994.
PF   31-MAR-1994; G00694.
PR   07-APR-1993; GB-007250.
PR   10-FEB-1994; GB-002573.
PA   (CANC-) CANCER RES INST.
PA   (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.
PI   Ashworth A,  Hughes DA,  Marshall CJ;
DR   WPI; 94-333199/41.
PT   Screening for a substance which is an inhibitor of mammalian MAP
PT   kinase - and is used in the mfr. of a medicament for the
PT   treatment of e.g. an inflammatory disorder or neurological
PT   disease.
PS   Disclosure; Fig. 7a; 53pp; English.
CC   STY2, STY3, STY4 and STY5 cDNAs (given in Q72858-61), encoding
CC   peptides having MAP-kinase-phosphatase activity, were derived from
CC   human squamous A431 cell polyA+ mRNAs.  Residues 177-255 of the
CC   encoded STY2-5 peptides were deduced (R63593-96) and demonstrated
CC   to show homology with the corresponding region of the previously
CC   isolated human CL100 gene product (R63592).
SQ   Sequence   79 AA;

 R63592  Length: 79  February 11, 2000 15:48  Type: P  Check: 2807  ..

       1  ILPFLYLGSA YHASRKDMLD ALGITALINV SANCPNHFEG HYQYKSIPVE

      51  DNHKADISSW FNEAIDFIDS IKNAGGRVF

!!AA_SEQUENCE 1.0
ID   R63602 standard; Peptide; 367 AA.
AC   R63602;
DT   09-MAY-1995  (first entry)
DE   MAP-kinase-phosphatase CL100.
KW   MAP-kinase inhibitor; MAPK; MAP-kinase-phosphatase; CL100; STY8;
KW   neurological disease; proliferative disorder; inflammation;
KW   cardiovascular disease.
OS   Homo sapiens.
PN   WO9423039-A.
PD   13-OCT-1994.
PF   31-MAR-1994; G00694.
PR   07-APR-1993; GB-007250.
PR   10-FEB-1994; GB-002573.
PA   (CANC-) CANCER RES INST.
PA   (CANC-) INST CANCER RES ROYAL CANCER HOSPITAL.
PI   Ashworth A,  Hughes DA,  Marshall CJ;
DR   WPI; 94-333199/41.
PT   Screening for a substance which is an inhibitor of mammalian MAP
PT   kinase - and is used in the mfr. of a medicament for the
PT   treatment of e.g. an inflammatory disorder or neurological
PT   disease.
PS   Disclosure; Fig. 7d; 53pp; English.
CC   STY8 cDNA (given in Q72864), encoding a peptide having
CC   MAP-kinase-phosphatase activity, was derived from a human
CC   brain cDNA library.  The STY8 protein (R63601) encoded by this
CC   full-length cDNA showed homology to the previously isolated
CC   human CL100 gene product (R63602).
SQ   Sequence   367 AA;

 R63602  Length: 367  February 11, 2000 15:48  Type: P  Check: 5661  ..

       1  MVMEVGTLDA GGLRALLGER AAQCLLLDCR SFFAFNAGHI AGSVNVRFST

      51  IVRRRAKGAM GLEHIVPNAE LRGRLLAGAY HAVVLLDERS AALDGAKRDG

     101  TLALAAFALC REARAAQVFF LKGGYEAFSA SCPELCSKQS TPMGLSLPLS

     151  TSVPDSAESG CSSCSTPLYD QGGPVEILPF LYLGSAYHAS RKDMLDALGI

     201  TALINVSANC PNHFEGHYQY KSIPVEDNHK ADISSNFMEA IDFIDSIKNA

     251  GGRVFVHCQA GISRSATICL AYLMRTNRVK LDEAFEFVKQ RRSIISPNFS

     301  FMGQLLQFES QVLPHSCSAE AGSPAMAVLD RGTSTTTVFN FPVSIPVHST

     351  NSALSYLHSP ITTSPSC

!!AA_SEQUENCE 1.0
ID   R58635 standard; protein; 124 AA.
AC   R58635;
DT   15-MAY-1995  (first entry)
DE   Amylase inhibitor protein 0.26 AI.
KW   Amylase inhibitor; wheat; cation exchange; inhibitor; amylase; human;
KW   pancreatic; alpha-amylase; homodimer; blood glucose level;
KW   insulin secretion; appetite; food additive.
OS   Triticum durum.
PN   EP-618229-A.
PD   05-OCT-1994.
PF   24-MAR-1994; 302148.
PR   29-MAR-1993; JP-091881.
PR   28-MAY-1993; JP-148423.
PA   (NAGA-) NAGATA SANGYO CO LTD.
PA   (NISS ) NISSHIN FLOUR MILLING CO.
PI   Matsubara H,  Miyazaki T,  Morimoto T,  Murayama R;
DR   WPI; 94-304407/38.
PT   New amylase inhibitor and process of prepn. from wheat - is
PT   useful for inhibiting increase in blood glucose level,
PT   controlling insulin secretion, suppressing appetite and as a food
PT   additive
PS   Claim 1; Page 15; 21pp; English.
CC   The amino acid sequence of a novel protein from the amylase
CC   inhibitor (AI) family.  By size and electrophoretic measurements,
CC   this protein has been called 0.26 AI.  The protein was isolated from
CC   wheat extracts by a novel process comprising (i) extracting a
CC   wheat-based material to produce a solution containing the AI, (ii)
CC   optionally purifying the solution to remove contaminants using a cation
CC   exchanger to absorb the AI, (iii) treating this exchanger with alkali
CC   at pH 9-13 to elute the AI, (iv) immediately adjusting the pH of the
CC   elute within neutral or acid range and (v) recovering the AI from this
CC   solution.  The novel AI has a very high inhibitory activity against an
CC   amylase especially human pancreatic alpha-amylase.  The activity of
CC   the AI is 5-30 times higher than other amylases of wheat origin such as
CC   0.53 AI or 0.28 AI.  The protein occurs as a homodimer with subunits of
CC   molecular masses 12500.  The protein can be used in combination with
CC   another amylase inhibitor (0.19) (R58636) as an agent for inhibiting
CC   increase in blood glucose level, controlling insulin secretion or
CC   suppressing appetite, and as a food additive, the total content of both
CC   proteins in the material being at least 20% by weight.
SQ   Sequence   124 AA;

 R58635  Length: 124  February 11, 2000 15:48  Type: P  Check: 2685  ..

       1  SGPWMCYPGY AFKVPALPGC RPVLKLQCNG SQVPEAVLRD CCQQLADISE

      51  WCRCGALYSM LDSMYKEHGV QEGQAGTGAF PSCRREVVKL TAASITAVCK

     101  LPIVIDASGD GAYVCKGVAA YPDA

!!AA_SEQUENCE 1.0
ID   R58636 standard; protein; 124 AA.
AC   R58636;
```

DT 15-MAY-1995 (first entry)
DE Amylase inhibitor protein 0.19 AI.
KW Amylase inhibitor; wheat; cation exchange; inhibitor; amylase; human;
KW pancreatic; alpha-amylase; homodimer; blood glucose level;
KW insulin secretion; appetite; food additive.
OS Triticum durum.
PN EP-618229-A.
PD 05-OCT-1994.
PF 24-MAR-1994; 302148.
PR 29-MAR-1993; JP-091881.
PR 28-MAY-1993; JP-148423.
PA (NAGA-) NAGATA SANGYO CO LTD.
PA (NISS) NISSHIN FLOUR MILLING CO.
PI Matsubara H, Miyazaki T, Morimoto T, Murayama R;
DR WPI; 94-304407/38.
PT New amylase inhibitor and process of prepn. from wheat - is
PT useful for inhibiting increase in blood glucose level,
PT controlling insulin secretion, suppressing appetite and as a food
PT additive
PS Claim 6; Page 16; 21pp; English.
CC The amino acid sequence of a protein from the amylase inhibitor (AI)
CC family. By size and electrophoretic measurements, this protein is called
CC 0.19 AI. The protein occurs as a homodimer, the subunits having the amino
CC acid sequence given. This protein can be used in combination with a
CC novel amylase inhibitor (0.26) (R58635) as an agent for inhibiting
CC increase in blood glucose level, controlling insulin secretion or
CC suppressing appetite, and as a food additive, the total content of both
CC proteins in the material being at least 20% by weight. The novel protein
CC was isolated from wheat extracts by a novel process comprising
CC (i) extracting a wheat-based material to produce a solution containing
CC the AI, (ii) optionally purifying the solution to remove contaminants
CC using a cation exchanger to absorb the AI, (iii) treating this exchanger
CC with alkali at pH 9-13 to elute the AI, (iv) immediately adjusting the pH
CC of the elute within neutral or acid range and (v) recovering the AI from
CC this solution. The novel AI has a very high inhibitory activity against
CC an amylase especially human pancreatic alpha-amylase. The activity of
CC the AI is 5-30 times higher than other amylases of wheat origin such as
CC 0.53 AI or 0.28 AI.
SQ Sequence 124 AA;

R58636 Length: 124 February 11, 2000 15:48 Type: P Check: 3320 ..

1 SGPWMCYPGQ AFQVPALPAC RPLLRLQCNG SQVPEAVLRD CCQQLAHISE

51 WCRCGALYSM LDSMYKEHGA QEGQAGTGAF PRCRREVVKL TAASITAVCR

101 LPIVVDASGD GAYVCKDVAA YPDA

!!AA_SEQUENCE 1.0
ID R58587 standard; Protein; 246 AA.
AC R58587;
DT 09-MAY-1995 (first entry)
DE Nicotineamide adenine dinucleotide synthetase N-terminal fragment.
KW Nicotineamide adenine dinucleotide; NAD; synthetase; NADS; E.coli.
OS Bacillus stearothermophilus.
FH Key Location/Qualifiers
FT peptide 2. .245
FT /label= mature peptide
PN J06225771-A.
PD 16-AUG-1994.
PF 01-FEB-1993; 014650.
PR 01-FEB-1993; JP-014650.
PA (ASAH) ASAHI KASEI KOGYO KK.
DR WPI; 94-298801/37.
DR N-PSDB; Q70537.
PT DNA coding NAD synthetase - used for the prepn. of NAD synthetase
PS Claim 2; Column 17-20; 16pp; Japanese.
CC The amino acid sequence for the N-terminal region of nicotineamide
CC adenine dinucleotide (NAD) synthetase (NADS). The gene encodes a protein
CC of 245 amino acids.The gene was inserted into E.coli in an expression
CC plasmid. This was used to produce the NADS protein. This method allows
CC the efficient preparation of NADS polypeptide.
SQ Sequence 246 AA;

R58587 Length: 246 February 11, 2000 15:48 Type: P Check: 9708 ..

1 MQEKIDKLVQ WLRDQVSSAG LNGAVVGISG GIDSAVVAHL IKRAFPDDSL

51 GLIMPCKSNP KDMEDALKVV KSCGIRHLVI DLTEAHRTLF GAVEAELKAI

101 GEWSEERARL GDANTRARLR MTTLYAVANN YGYLVVGTDN AAEWHTGYFT

151 KYGDGGVDLV PLIHFTKGEV REMGRLLGVP EEIIKKAPSA GLWEGQTDES

201 EMGTTYEMID KYLKGEEIPE RDRKIIERLH ERSHHKRQLA IAPPKF

!!AA_SEQUENCE 1.0
ID R63443 standard; Protein; 120 AA.
AC R63443;
DT 06-JUN-1995 (first entry)
DE Amino terminal sequence of the rubisco activase - AGP small subunit
DE fusion enzyme.
KW Barley; endosperm; ADP glucose pyrophosphorylase; AGP; starch;
KW enzyme; transgenic plant; transit peptide; spinach; pPPS4; ss.
OS Synthetic.
PN WO9424292-A.
PD 27-OCT-1994.
PF 07-APR-1994; E01082.
PR 08-APR-1993; GB-007408.
PA (DANI-) DANISCO AS.
PI Kleczkowski L, Marcussen J, Okkels F, Olsen O, Poulsen P;
PI Villand P;
DR WPI; 94-341871/42.
PT Transgenic starch producing organism - comprises exogenous ADP
PT glucose pyrophosphorylase enzyme DNA sequence
PS Example; Fig 8; 88pp; English.
CC Plasmid pPPS4 is a pVictorIV deriv. in which a 3 kb KpnI fragment
CC contg. the construct:
CC patatin promoter - spinach rubisco activase transit peptide -
CC small subunit AGP from barley endosperm - 35S terminator;
CC is inserted in the KpnI site. R63443 gives the N-terminal sequence
CC of the rubisco activase-AGP small subunit fusion enzyme. pPPS4
CC is claimed.
SQ Sequence 120 AA;

R63443 Length: 120 February 11, 2000 15:48 Type: P Check: 8722 ..

1 MATAVSTVGA ATRAPLNLNG SSAGASVPTS GFLGSSLKKH TNVRFPSSSR

51 TTSMTVKAAE NEEKNTDKWA HLAKDFSDDQ LDIRRGKGMV DSLGSMDVPL

101 ASKVPLPSPS KHEQCNVYSH

!!AA_SEQUENCE 1.0
ID R63444 standard; Protein; 120 AA.
AC R63444;
DT 06-JUN-1995 (first entry)
DE Amino terminal sequence of the rubisco activase - AGP large subunit
DE fusion enzyme.
KW Barley; endosperm; ADP glucose pyrophosphorylase; AGP; starch;
KW enzyme; transgenic plant; transit peptide; spinach; pPPL4; ss.
OS Synthetic.
PN WO9424292-A.
PD 27-OCT-1994.
PF 07-APR-1994; E01082.
PR 08-APR-1993; GB-007408.
PA (DANI-) DANISCO AS.
PI Kleczkowski L, Marcussen J, Okkels F, Olsen O, Poulsen P;
PI Villand P;
DR WPI; 94-341871/42.
PT Transgenic starch producing organism - comprises exogenous ADP
PT glucose pyrophosphorylase enzyme DNA sequence
PS Example; Fig 8; 88pp; English.
CC Plasmid pPPL4 is a pVictorIV SGiN Man deriv. in which a 3.2 kb
CC EcoRI fragment contg. the construct:
CC patatin promoter - spinach rubisco activase transit peptide -

CC large subunit AGP from barley endosperm - 35S terminator;
CC is inserted in the EcoRI site. R63444 gives the N-terminal sequence
CC of the rubisco activase-AGP large subunit fusion enzyme. pPPL4
CC is claimed.
SQ Sequence 120 AA;

R63444 Length: 120 February 11, 2000 15:48 Type: P Check: 7811 ..

1 MATAVSTVGA ATRAPLNLNG SSAGASVPTS GFLGSSLKKH TNVRFPSSSR

51 TTSMTVKAAE NEEKNTDKWA HLAKDFSDDQ LDIRRGKGMV DSLGIHMQFS

101 SVLPLEGKAC VSPVRREGSA

!!AA_SEQUENCE 1.0
ID R62507 standard; Protein; 313 AA.
AC R62507;
DT 26-JUN-1995 (first entry)
DE Galactosyl transferase 3' clone product.
KW Gal-alpha (1,3) galactosyl transferase; xenograft; transplant;
KW rejection.
OS Sus scrofa domestica.
PN WO9421799-A.
PD 29-SEP-1994.
PF 15-MAR-1994; AU0126.
PR 16-MAR-1993; AU-007854.
PA (AUST-) AUSTIN RES INST.
PI Mckenzie IFC, Sandrin MS;
DR WPI; 94-317019/39.
DR N-PSDB; Q74711.
PT DNA sequences encoding Gal-alpha (1,3)galactosyl transferase -
PT and clones contg. such sequences are used in xenograft therapies
PS Disclosure; Page 31; 50pp; English.
CC The sequence is that of the product of the porcine Gal-alpha (1,3)
CC galactosyl transferase gene which produces a Gal epitope on the
CC surface of porcine cells. This epitope is recognised by antibodies
CC which are responsible for hyperacute rejection of xenotransplanted
CC pig cells, tissues and organs.
CC See also R62508.
SQ Sequence 313 AA;

R62507 Length: 313 February 11, 2000 15:48 Type: P Check: 2303 ..

1 VPSSNSASQS PQAMTDPCSP RLSYLSKAIL TLCFVTRKPP EVVTITRWKA

51 PVVWEGTYNR AVLDNYYAKQ KITVGLTVFA VGRYIEHYLE EFLISANTYF

101 MVGHKVIFYI MVDDISRMPL IELGPLRSFK VFEIKSEKRW QDISMMRMKT

151 IGEHILAHIQ HEVDFLFCID VDQVFQNNFG VETLGQSVAQ LQAWWYKAHP

201 DEFTYERPKE SAAYIPFRQG DFYYHAAILG GTPTQVLNIT QECFKGILQD

251 KENDIEAEWH DESGLNKYFL LNKPTKILSP EYCWDYHIGM SVDIRIVKGA

301 WQKKEYNLVR NNI

!!AA_SEQUENCE 1.0
ID R66355 standard; peptide; 216 AA.
AC R66355;
DT 28-JUN-1995 (first entry)
DE Ii protein.
KW Ii protein; immunomodulator; antigen presentation; MHC;
KW major histocompatibility complex.
OS Homo sapiens.
PN WO9426773-A.
PD 24-NOV-1994.
PF 18-MAY-1994; U05617.
PR 19-MAY-1993; US-064400.
PA (ANTI-) ANTIGEN EXPRESS INC.
PI Humphreys RE;
DR WPI; 95-006697/01.
PT Identification of mutant Ii mols having altered endoprotease
PT cleavage sites - and useful for identifying immunomodulators,
PT esp. peptide(s) which enhance or inhibit antigen presentation
PS Disclosure; Page 38-39; 50pp; English.
CC Immunomodulatory peptides of Ii protein (full sequence given
CC R66355) have been identified. The Ii peptide given in R66352
CC enhances MHC class II-mediated antigen presentation, while the
CC Ii peptides given in R66353-54 inhibit antigen presentation.
SQ Sequence 216 AA;

R66355 Length: 216 February 11, 2000 15:48 Type: P Check: 1046 ..

1 MDDQRDLISN NEQLPMLGRR PGAPESKCSR GALYTGFSIL VTLLLAGQAT

51 TAYFLYQQQG RLDKLTVTSQ NLQLENLRMK LPKPPKPVSK MRMATPLLMQ

101 ALPMGALPQG PMQNATKYGN MTEDHVMHLL QNADPLKVYP PLKGSFPENL

151 RHLKNTMETL DWKVFESWMH HWLLFEMSRH SLEQKPTDAP PKESLELEDP

201 SSGLGVTKQD LGPVPM

!!AA_SEQUENCE 1.0
ID R63159 standard; Protein; 376 AA.
AC R63159;
DT 23-JUN-1995 (first entry)
DE Mouse growth differentiation factor-8 protein.
KW Growth differentiation factor-8; GDF-8; cell proliferation;
KW adipocyte; obesity; transforming growth factor-beta.
OS Mus musculus.
PN WO9421681-A.
PD 29-SEP-1994.
PF 18-MAR-1994; U03019.
PR 19-MAR-1993; US-033923.
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MED.
PI Lee S, Mcpherron AC;
DR WPI; 94-316943/39.
DR Q-PSDB; Q76371.
PT New growth differentiation factor 8 - useful for treatment and
PT diagnosis of cell proliferative disorders esp. of muscle.
PS Claim 3; Page 47; 84pp; English.
CC GDF-8 can be used to maintain cells before transplantation; to
CC improve efficiency of cell fusion and to treat obesity or diseases
CC related to abnormal adipocyte proliferation.
SQ Sequence 376 AA;

R63159 Length: 376 February 11, 2000 15:48 Type: P Check: 2293 ..

1 MMQKLQMYVY IYLFMLIAAG PVDLNEGSER EENVEKEGLC NACAWRQNTR

51 YSRIEAIKIQ ILSKLRLETA PNISKDAIRQ LLPRAPPLRE LIDQYDVQRD

101 DSSDGSLEDD DYHATTETII TMPTESDFLM QADGKPKCCF FKFSSKIQYN

151 KVVKAQLWIY LRPVKTPTTV FVQILRLIKP MKDGTRYTGI RSLKLDMSPG

201 TGIWQSIDVK TVLQNWLKQP ESNLGIEIKA LDENGHDLAV TFPGPGEDGL

251 NPFLEVKVTD TPKRSRRDFG LDCDEHSTES RCCRYPLTVD FEAFGWDWII

301 APKRYKANYC SGECEFVFLQ KYPHTHLVHQ ANPRGSAGPC CTPTKMSPIN

351 MLYFNGKEQI IYGKIPAMVV DRCGCS

!!AA_SEQUENCE 1.0
ID R63160 standard; Protein; 375 AA.
AC R63160;
DT 23-JUN-1995 (first entry)
DE Human growth differentiation factor-8 protein.
KW Growth differentiation factor-8; GDF-8; cell proliferation;
KW adipocyte; obesity; transforming growth factor-beta.
OS Homo sapiens.
PN WO9421681-A.
PD 29-SEP-1994.

PF 18-MAR-1994; U03019.
PR 19-MAR-1993; US-033923.
PA (UYJO) UNIV JOHNS HOPKINS SCHOOL MED.
PI Lee S, Mcpherron AC;
DR WPI; 94-316943/39.
DR Q-PSDB; Q76372.
PT New growth differentiation factor 8 - useful for treatment and
PT diagnosis of cell proliferative disorders esp. of muscle.
PS Claim 3; Page 58; 84pp; English.
CC GDF-8 can be used to maintain cells before transplantation; to
CC improve efficiency of cell fusion and to treat obesity or diseases
CC related to abnormal adipocyte proliferation.
SQ Sequence 375 AA;

R63160 Length: 375 February 11, 2000 15:48 Type: P Check: 1814 ..

1 MQKLQLCVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACTWRQNTKS

51 SRIEAIKIQI LSKLRLETAP NISKDVIRQL LPKAPPLREL IDQYDVQRDD

101 SSDGSLEDDD YHATTETIIT MPTESDFLMQ VDGKPKCCFF KFSSKIQYNK

151 VVKAQLWIYL RPVETPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

201 GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPGPGEDGLN

251 PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

301 PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

351 LYFNGKEQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID R66665 standard; peptide; 25 AA.
AC R66665;
DT 21-JUL-1995 (first entry)
DE Binding peptide B (84 TS)
KW Inner capsid protein; VP6; carrier; vaccine; rotavirus; VP3;
KW binding peptide.
OS Rotavirus.
PN US5374426-A.
PD 20-DEC-1994.
PF 03-SEP-1986; 903222.
PR 03-SEP-1986; US-903222.
PR 02-SEP-1987; US-092120.
PR 30-OCT-1989; US-429147.
PA (UYSA-) UNIV SASKATCHEWAN.
PI Frenchick PJ, Mullin-ready KF, Sabara MI;
DR WPI; 95-035623/05.
PT Immunogenic complex contg. rotavirus VP6 protein as carrier -
PT coupled, esp. by protein-protein interaction, to an epitope
PT contg. molecule.
PS Claim 11(b); Column 26; 23pp; English.
CC An immunological carrier complex comprises an epitope-bearing
CC molecule bound to the rotavirus inner capsid protein VP6 (given in
CC R66669) using a binding peptide such as those given in R66664-8.
CC Peptide B (or 84 TS) (R66665) is derived from rotavirus VP3
CC (231-254).
SQ Sequence 25 AA;

R66665 Length: 25 February 11, 2000 15:48 Type: P Check: 4995 ..

1 CNIAPASIVS RNIVYTRAQP NQDIA

!!AA_SEQUENCE 1.0
ID R66680 standard; peptide; 25 AA.
AC R66680;
DT 21-JUL-1995 (first entry)
DE Peptide 84 TS-CYS.
KW Inner capsid protein; VP6; carrier; vaccine; rotavirus; VP3;
KW binding peptide.
OS Synthetic.
PN US5374426-A.
PD 20-DEC-1994.
PF 03-SEP-1986; 903222.
PR 03-SEP-1986; US-903222.
PR 02-SEP-1987; US-092120.
PR 30-OCT-1989; US-429147.
PA (UYSA-) UNIV SASKATCHEWAN.
PI Frenchick PJ, Mullin-ready KF, Sabara MI;
DR WPI; 95-035623/05.
PT Immunogenic complex contg. rotavirus VP6 protein as carrier -
PT coupled, esp. by protein-protein interaction, to an epitope
PT contg. molecule.
PS Example 2; Column 18; 23pp; English.
CC An immunological carrier complex comprises an epitope-bearing
CC molecule bound to the rotavirus inner capsid protein VP6 (given in
CC R66669) using a binding peptide such as peptide B (84 TS)
CC (R66665). Variants of peptide B were also tested for binding to VP6.
CC Only those that generated a spatial arrangement of Cys and Arg
CC residues in a suitable 3-dimensional orientation (e.g. R66666-68)
CC were suitable. 84 TS-CYS (R66680) did not bind to VP6.
SQ Sequence 25 AA;

R66680 Length: 25 February 11, 2000 15:48 Type: P Check: 5136 ..

1 NISPASIVSR NIVYTRAQPN QSIAC

!!AA_SEQUENCE 1.0
ID R66682 standard; peptide; 25 AA.
AC R66682;
DT 21-JUL-1995 (first entry)
DE Peptide MONOSER.
KW Inner capsid protein; VP6; carrier; vaccine; rotavirus; VP3;
KW binding peptide.
OS Synthetic.
PN US5374426-A.
PD 20-DEC-1994.
PF 03-SEP-1986; 903222.
PR 03-SEP-1986; US-903222.
PR 02-SEP-1987; US-092120.
PR 30-OCT-1989; US-429147.
PA (UYSA-) UNIV SASKATCHEWAN.
PI Frenchick PJ, Mullin-ready KF, Sabara MI;
DR WPI; 95-035623/05.
PT Immunogenic complex contg. rotavirus VP6 protein as carrier -
PT coupled, esp. by protein-protein interaction, to an epitope
PT contg. molecule.
PS Example 2; Column 18; 23pp; English.
CC An immunological carrier complex comprises an epitope-bearing
CC molecule bound to the rotavirus inner capsid protein VP6 (given in
CC R66669) using a binding peptide such as peptide B (84 TS)
CC (R66665). Variants of peptide B were also tested for binding to VP6.
CC Only those that generated a spatial arrangement of Cys and Arg
CC residues in a suitable 3-dimensional orientation (e.g. R66666-68)
CC were suitable. MONOSER (R66682) did not bind to VP6.
SQ Sequence 25 AA;

R66682 Length: 25 February 11, 2000 15:48 Type: P Check: 5012 ..

1 CNIAPASIVS RNIVYTSAQP NQDIA

!!AA_SEQUENCE 1.0
ID R67596 standard; Protein; 331 AA.
AC R67596;
DT 20-JUL-1995 (first entry)
DE A. aculeatus pectin methylesterase.
KW Pectin methylesterse; PME; strain CBS 101.43; demethylation; pectin;
KW food industry; fruit; vegetable; processing; plant cell wall;
KW modification; pectin lyase; glucanase; firmness.
OS Aspergillus aculeatus.
PN WO9425575-A.
PD 10-NOV-1994.
PF 28-APR-1994; DK0173.
PR 30-APR-1993; DK-000487.
PR 28-OCT-1993; DK-001217.

PA (NOVO) NOVO-NORDISK AS.
PI Andersen LN, Budolfsen G, Christgau S, Dalboege H;
PI Heldt-Hansen HP, Kauppinen S, Kofod LV;
DR WPI; 95-022253/03.
DR N-PSDB; Q79154.
PT Enzyme with pectin methylesterase activity - useful for
PT demethylation of pectin, e.g. for application in food industry.
PS Claim 1; Page 37; 53pp; English.
CC This sequence represents an enzyme showing pectin methylesterse (PME)
CC activity. The DNA sequence encoding this protein was isolated from
CC strain CBS 101.43. The PME enzyme is useful for demethylation of
CC pectin. It can therefore be used in the food industry for fruit and
CC vegetable processing. It modifies plant cell walls and can be used
CC with other plant cell wall degrading enzymes, eg. pectin lyase,
CC glucanase, etc. It can be used to improve the firmness of a pectin-
CC contg. material or to increase viscosity of a pectin-contg. material.
CC The enzyme has an optimum pH if 4.5, a temperature optimum of 45 deg
CC C and a molecular weight of 43 kD.
SQ Sequence 331 AA;

R67596 Length: 331 February 11, 2000 15:48 Type: P Check: 5465 ..

1 MVKSVLASAT FAVSATAASR TTAPSGAIVV AKSGGDYTTI GDAIDALSTS

51 TTDTQTIFIE EGTYDEQVYL PAMTGKVIIY GQTENTDSYA DNTVTITHAI

101 SYEDAGESDD TTATFRNKAV GSQVYNTNIA NTCGQACHQA LALSAWADQQ

151 GYYGCNFTGY QDTTTAQTGN QTYINSYIEG AVDFIFGQHA RAWFQNVDIR

201 VVEGPTSASI TANGRSSETD TSYYVINKST VAAKEGDDVA EGTYYTGRPW

251 SEYARVVFQQ TSMTNVINST GWTEWSTSTP NTEYVTFGEY ANQPAPAPRA

301 PAHVAEKTDA KTTITDITGS DYTSWVDTSY F

!!AA_SEQUENCE 1.0
ID R71324 standard; Protein; 329 AA.
AC R71324;
DT 21-OCT-1995 (first entry)
DE Acetyl-CoA-reductase.
KW Acetyl-CoA-reductase; transgenic plant;
KW poly-beta-hydroxyalkanoate; poly-beta-hydroxybutyrate;
KW biodegradable thermoplastic.
OS Alcaligenes eutrophus.
PN WO9505472-A.
PD 23-FEB-1995.
PF 17-AUG-1994; U09265.
PR 17-AUG-1993; US-108193.
PR 06-JUN-1994; US-254357.
PA (UNMS) UNIV MICHIGAN STATE.
PI Nawrath C, Poirier Y, Somerville CR;
DR WPI; 95-098770/13.
DR N-PSDB; Q85642.
PT Transgenic plant material with plastid(s) contg. the enzymes for
PT synthesis of poly:hydroxy:alkanoate(s) - express
PT poly:hydroxy:butyrate and have good growth and seed formation.
PS Claim 2; Page 60-61; 88pp; English.
CC The acetyl-CoA-reductase gene (phbB) from A. eutrophus is cloned
CC under the control of an Arabidopsis thaliana seed storage protein
CC promoter for plastid tissue-specific gene expression in a
CC transgenic plant. When expressed with the 3-ketothiolase
CC (phbA) and poly-beta-hydroxyalkanoate-synthase (phbC) genes, a
CC poly-beta-hydroxyalkanoate (PHA), specifically poly-beta-
CC hydroxybutyrate (PHB), is expressed in the transgenic plant
CC (preferably a Brassica e.g. rape). PHB and related PHAs are
CC biodegradable thermoplastics with many useful applications.
SQ Sequence 329 AA;

R71324 Length: 329 February 11, 2000 15:48 Type: P Check: 7564 ..

1 MASMISSSAV TTVSRASRGQ SAAVAPFGGL KSMTGFPVKK VNTDITSITS

51 NGGRVKCMQV WPPIGKKKFE TLSYLPPLTR DSRVTQRIAY VTGGMGGIGT

101 AICQRLAKDG FRVVAGCGPN SPRREKWLEQ QKALGFDFIA SEGNVADWDS

151 TKTAFDKVKS EVGEVDVLIN NAGITRDVVF RKMTRADWDA VIDTNLTSLF

201 NVTKQVIDGM ADRGWGRIVN ISSVNGQKGQ FGQTNYSTAK AGLHGFTMAL

251 AQEVATKGVT VNTVSPGYIA TDMVKAIRQD VLDKIVATIP VKRLGLPEEI

301 ASICAWLSSE ESGFSTGADF SLNGGLHMG

!!AA_SEQUENCE 1.0
ID R70100 standard; peptide; 274 AA.
AC R70100;
DT 11-DEC-1995 (first entry)
DE Lettuce infectious yellows virus open reading frame 3 protein.
KW Expression vector; expression cassette; Agrobacterium tumefaciens; plant;
KW lettuce infectious yellows virus coat protein; amplify; PCR; primer;
KW heat shock protein.
OS Synthetic.
PN WO9502056-A.
PD 19-JAN-1995.
PF 13-JUN-1994; U06430.
PR 09-JUL-1993; US-090532.
PR 15-OCT-1993; US-138138.
PR 01-NOV-1993; US-146780.
PA (REGC) UNIV CALIFORNIA.
PA (UPJO) UPJOHN CO.
PI Boeshore ML, Carney KJ, Falk BW, Klaassen V;
DR WPI; 95-106554/14.
DR N-PSDB; Q83174.
PT New DNA encoding proteins of lettuce infectious yellows virus -
PT and related vectors, transformed cells and transgenic plants
PT resistant to virus infection
PS Claim 1; Fig 14; 92pp; English.
CC The amino acid sequence encoded by the open reading frame 3 of the
CC lettuce infectious yellows virus (LIYV) RNA 1. The gene was amplified by
CC PCR using the primers Q83175-6. The resultant product was used to
CC construct expression cassettes expressing the gene as a sense (pEPG170)
CC message. The vectors containing the cassettes can be transformed into
CC Agrobacterium tumefaciens for transfer of the genes into plants. Plants
CC containing the sense or antisense vectors are resistant to LIYV
CC infections.
SQ Sequence 274 AA;

R70100 Length: 274 February 11, 2000 15:48 Type: P Check: 4652 ..

1 MIMMSPLYAL TKQCVIDTAY RLAVPTQHCA IYTVACRILF LSVGFMTIVK

51 LCGFKMDTSS FIASIEKDNL MDCLISLVEM RDRLRLCNDF PILNYGVNIL

101 ELLIGKRLNK INNLKNCYVI RELITINISK EWVGKQALKV GLHCFLNLSQ

151 AESRHVKYLL SDKESLNKMN FSRYYVPKVV TDLYLDLIGV LYVNTGYNID

201 LVEKFIFDKL EFLVYDGEEG FKSPQVEYND ICTVYNLKPI IKYNRWHTDG

251 SIVIECGDVI GKGINKTKKK ICNK

!!AA_SEQUENCE 1.0
ID R72715 standard; Protein; 359 AA.
AC R72715;
DT 02-DEC-1995 (first entry)
DE hisC gene product of Salmonella typhimurium.
KW hisG; hisC; Salmonella typhimurium; mutant; auxotroph; detection;
KW identification; mutagen; transition; transversion.
OS Salmonella typhimurium.
PN WO9510629-A.
PD 20-APR-1995.
PF 14-OCT-1994; U11617.
PR 15-OCT-1993; US-137627.
PA (REGC) UNIV CALIFORNIA.

```
PI   Ames BN,  Gee P,  Maron DM;
DR   WPI; 95-161813/21.
DR   N-PSDB; Q89706.
PT   Salmonella auxotrophic mutants for detecting mutagenic agents -
PT   comprise point mutation(s) in histidine genes hisG and hisC ; for
PT   identifying specific base changes induced by a mutagen
PS   Disclosure; Page 96-97; 135pp; English.
CC   Salmonella bacteria having substitutions in DNA sequences which
CC   render the bacteria non-selectable for a known Salmonella
CC   characteristic may be used for characterising specific DNA base
CC   substitutions mutations induced by a mutagen causing a reversion
CC   back to the wild type phenotype.  The mutant Salmonella are
CC   auxotrophic and only revert back to the wild type phenotype if the
CC   reverse mutation occurs at a specific site.  Base transitions/
CC   transversions are identified by exposing 2-4 derivative strains to a
CC   mutagen.  The strains are then cultured in medium selecting for a
CC   specific characteristic and the base transition/transversions are
CC   identified. His mutants were sequenced to identify the type and
CC   physical location of the mutations responsible for the his-
CC   phenotype.
SQ   Sequence   359 AA;

 R72715  Length: 359  February 11, 2000 15:48  Type: P  Check: 516  ..

       1  MSTENTLSVA DLARENVRNL VPYQSARRLG GNGDVWLNAN EFPTAVEFQL

      51  TQQTLNRYPE CQPKAVIENY AQYAGVKPEQ VLVSRGADEG IELVIRAFCE

     101  PGKDAILYCP PTYGMYSVSA ETIGVERRTV PALENWQLDL QGISDNLDGT

     151  KVVFVCSPNN PTGQLINPQD LRTLLELTRG KAIVVADEAY IEFCPQATLT

     201  GWLVEYPHLV ILRTLSKAFA LAGLRCGFTL ANEEVINLLL KVIAPYPLST

     251  PVADIAAQAL CPQGINAMRD RVAQTVQERQ YLVNALQQTA CVEHVFDSET

     301  NYILARFTAS SSVFKSLWDQ GIILRDQNKQ PSLSGCLRIT VGTRQENQRV

     351  IDALRAEPV

!!AA_SEQUENCE 1.0
ID   R74180 standard; protein; 248 AA.
AC   R74180;
DT   01-JAN-1996  (first entry)
DE   Type I ribosome-inactivating protein luffin.
KW   Ribosome inactivating protein; RIP; Type I; cytotoxin; immunotoxin.
OS   Plant.
FH   Key             Location/Qualifiers
FT   misc_difference 14
FT                   /label= invariant residue
FT                   /note= "in ricin A-chain and the Type I RIPs"
FT   misc_difference 22
FT                   /label= see above
FT   misc_difference 70
FT                   /label= see above
FT   misc_difference 110
FT                   /label= see above
FT   misc_difference 131
FT                   /label= see above
FT   misc_difference 157. .158
FT                   /label= see above
FT   misc_difference 160
FT                   /label= see above
FT   misc_difference 192
FT                   /label= see above
PN   US5416202-A.
PD   16-MAY-1995.
PF   09-DEC-1992; 988430.
PR   04-NOV-1991; US-787567.
PR   09-DEC-1992; US-988430.
PA   (XOMA ) XOMA CORP.
PI   Bernhard SL, Better MD, Carroll SF, Lane JA, Lei SP;
DR   WPI; 95-193480/25.
PT   Polynucleotide(s) encoding gelonin analogues - having a cystein
PT   residue for intermolecular bonding for the prodn. of immuno-toxins(s)
PS   Disclosure; Figure 4; 66pp; English.
CC   Analogues of Type I RIP are defined as non-naturally occurring
CC   polypeptides that share the ribosome-inactivating activity of the
CC   natural protein but differ in AA sequence. Preferred analogues have
CC   a Cys available for disulfide bonding located at a posn. it its AA
CC   sequence from the posn. corresp. to posn. 251 in ricin A-chain RTA
CC   to the carboxy terminus of the analogue. (R74176 is the sequence of
CC   ricin A-chain RTA, which is a Type II RIP). The primary AA
CC   sequence of the Type I RIPs gelonin, BRIP, mormordin II, luffin
CC   [see Islam et al. Agricultural Biological Chem., 54(5) 1343-45 1991],
CC   alphatrichosanthin ([see Chow et al., J. Biol. Chem., 265, 8670-74
CC   1990], momordin I [see Ho et al., BBA, 1088, 311-14 1991],
CC   Mirabilis antiviral protein [see Habuka et al., J. Biol. Chem.,
CC   264(12) 6629-37 1989], pokeweed antiviral protein isolated from
CC   seeds [see Kung et al., Agric. Biol. Chem., 54(12), 3301-18 1990]
CC   and saporin [see Banatti et al., Eur. J. Biochem., 183, 465-70
CC   1989] are individually aligned with the primary sequence of the
CC   ricin A-chain [see Halling et al., Nucleic Acids Res., 13,
CC   8019-8033 1985] respectively in Figures 1-9. The AAs invariant
CC   among the ricin A-chain and the Type I RIPs are indicated in FT.
SQ   Sequence   248 AA;

 R74180  Length: 248  February 11, 2000 15:48  Type: P  Check: 8163  ..

       1  DVRFSLSGSS STSYSKFIGD LRKALPSNGT VYNLTILLSS ASGASRYTLM

      51  TLSNYDGKAI TVAVDVSQLY IMGYLVNSTS YFFNESDAKL ASQYVFKGST

     101  IVTLPYSGNY EKLQTAAGKI REKIPLGFPA LDSALTTIFH YDSTAAAAAF

     151  LVILQTTAEA SRFKYIEGQI IERISKNQVP SLATISLENS LWSALSKQIQ

     201  LAQTNNGTFK TPVVITDDKQ QRVEITNVTS KVVTKNIQLL LNYKQNVA

!!AA_SEQUENCE 1.0
ID   R84103 standard; Protein; 303 AA.
AC   R84103;
DT   20-MAR-1996  (first entry)
DE   Equine herpesvirus (EHV1) unique short (US2) gene prod. .
KW   Equine herpesvirus 1; EHV1; unique short; US2; recombinant; vaccine;
KW   attenuated virus; diagnosis.
OS   Equine herpesvirus 1.
FH   Key             Location/Qualifiers
FT   region          123. .140
FT                   /note= "EHV-1, EHV-4, HSV-1, PRV, HSV-2, MDV, and
FT                   IBR consereved US2 region"
PN   WO9522607-A1.
PD   24-AUG-1995.
PF   16-FEB-1995; U02087.
PR   17-FEB-1994; US-198094.
PA   (SYTR ) SYNTRO CORP.
PI   Chiang CH,  Cochran MD;
DR   WPI; 95-302714/39.
DR   N-PSDB; T00531.
PT   Recombinant equine herpes viruses pref. contg. a deletion in a
PT   region not essential for replication - used in vaccines to protect
PT   horses from infection
PS   Example 11; Pages 86-87; 159pp; English.
CC   T00531 encodes R84103 the equine herpesvirus 1 (EHV) unique short 2
CC   (US2) gene prod. . A recombinant attenuated EHV can be produced by
CC   deleting the above US2 gene (a region not essential for replication)
CC   from the viral DNA. The attenuated virus can be used as a foetal safe
CC   vaccine to protect an equine against EHV, or in a test to determine
CC   if an equine has been vaccinated against or is infected with EHV.
SQ   Sequence   303 AA;

 R84103  Length: 303  February 11, 2000 15:48  Type: P  Check: 5723  ..

       1  MGVVLITVVT VVDRHKALPN SSIDVDGHLW EFLSRQCFVL ASEPLGIPIV

      51  VRSADLYRFS SSLLTLPKAC RPIVRTRGAT AIALDRNGVV YHEDRMGVSI
```

```
     101  EWLSVLSGYN HLNSSLIINQ PYHLWVLGAA DLCKPVFDLI PGPKRMVYAE

     151  IADEFHKSWQ PPFVCGKLFE TIPWTTVEHN HPLKLRAAGG EDTVVGECGF

     201  SKHSSNSLVR PPTVKRVIYA VVDPARLREI PAPGRPLPRR RPSEGGMRAP

     251  RRRSRAPAAA RSTAAAATPP RPGDPRAPAA RRAGDVTWME RLLWGVFGRT

     301  STR

!!AA_SEQUENCE 1.0
ID   R81323 standard; Protein; 123 AA.
AC   R81323;
DT   02-APR-1996  (first entry)
DE   Humanized VLA-4 antibody 21.6 heavy chain variable region, Ha.
KW   Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
KW   antibody engineering.
OS   Mus musculus.
PN   WO9519790-A1.
PD   27-JUL-1995.
PF   25-JAN-1995; U01219.
PR   25-JAN-1994; US-186269.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones TS,  Leger OJ,  Saldanha J;
DR   WPI; 95-269276/35.
PT   New humanised antibodies against VLA-4 - used for inhibiting
PT   leukocyte adhesion to endothelial cells, partic. for treating
PT   inflammatory disease.
PS   Claim 11; Page 69; 105pp; English.
CC   The sequence encodes the mouse antibody 21.6 heavy chain variable
CC   region, Ha, directed against leukocyte adhesion molecule VLA-4. Cloned
CC   cDNA sequences of mouse 21.6 VL and VH (Q99889 and Q99892) regions are
CC   linked to human constant regions in the construction of a humanized
CC   antibody against VLA-4.  The 5' and 3' ends of the mouse cDNAs are
CC   modified using PCR primers (See Q99895-98) and then subcloned into
CC   mammalian cell expression vectors containing human kappa or gamma-1
CC   constant regions.  In the humanized light chain, amino acids L45,
CC   L49, L58 and L69 in the human kappa LC VR framework are replaced
CC   by the amino acid present in the equivalent position of the mouse
CC   21.6 Ig L chain.  Plasmids encoding the chimeric antibodies are
CC   transfected into COS cells.  The humanized antibodies can be used
CC   for inhibiting adhesion of a leukocyte to an endothelial cell and
CC   for treating inflammatory diseases such as multiple sclerosis.  They
CC   can also be used in the treatment of stroke, cerebral traumas,
CC   meningitis or encephalitis.  The antibodies can also be used for
CC   detecting VLA-4, for affinity purification or for generating
CC   anti-idiotype antibodies.
SQ   Sequence   123 AA;

 R81323  Length: 123  February 11, 2000 15:48  Type: P  Check: 9504  ..

       1  QVQLVQSGAE VKKPGASVKV SCKASGFNIK DTYIHWVRQA PGQRLEWMGR

      51  IDPANGYTKY DPKFQGRVTI TADTSASTAY MELSSLRSED TAVYYCAREG

     101  YYGNYGVYAM DYWGQGTLVT VSS

!!AA_SEQUENCE 1.0
ID   R81327 standard; Protein; 140 AA.
AC   R81327;
DT   23-MAR-1996  (first entry)
DE   Mouse VLA-4 antibody 21.6 light heavy variable region.
KW   Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
KW   antibody engineering.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   peptide         1. .19
FT                   /note= "signal peptide"
FT   region          20. .49
FT                   /note= "framework region 1"
FT   region          50. .54
FT                   /note= "complementarity determining region 1"
FT   region          55. .68
FT                   /note= "framework region 2"
FT   region          69. .85
FT                   /note= "complementarity determining region 2"
FT   region          86. .117
FT                   /note= "framework region 3"
FT   region          118. .131
FT                   /note= "complementarity determining region 3"
FT   region          132. .140
FT                   /note= "framework region 4"
PN   WO9519790-A1.
PD   27-JUL-1995.
PF   25-JAN-1995; U01219.
PR   25-JAN-1994; US-186269.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones TS,  Leger OJ,  Saldanha J;
DR   WPI; 95-269276/35.
DR   N-PSDB; Q99892.
PT   New humanised antibodies against VLA-4 - used for inhibiting
PT   leukocyte adhesion to endothelial cells, partic. for treating
PT   inflammatory disease.
PS   Disclosure; Fig 2; 105pp; English.
CC   The sequence represents the mouse antibody 21.6 heavy chain variable
CC   region directed against leukocyte adhesion molecule VLA-4.  Cloned
CC   cDNA sequences of mouse 21.6 VH and VL (see Q99889) regions are
CC   linked to human constant regions in the construction of a humanized
CC   antibody against VLA-4.  The 5' and 3' ends of the mouse cDNAs are
CC   modified using PCR primers (See Q99895-98) and then subcloned into
CC   mammalian cell expression vectors containing human kappa or gamma-1
CC   constant regions.  In the humanized heavy chain, amino acids H27,
CC   H28, H29, H30, H44 nd H71 in the human HC VR framework are replaced
CC   by the amino acid present in the equivalent position of the mouse
CC   21.6 Ig H chain.  Plasmids encoding the chimeric antibodies are
CC   transfected into COS cells.  The humanized antibodies can be used
CC   to inhibit adhesion of a leukocyte to an endothelial cell and
CC   to treat inflammatory diseases such as multiple sclerosis.  They
CC   can also be used in the treatment of stroke, cerebral traumas,
CC   meningitis or encephalitis.  The antibodies can also be used for
CC   detecting VLA-4, for affinity purification or for generating
CC   anti-idiotype antibodies.
SQ   Sequence   140 AA;

 R81327  Length: 140  February 11, 2000 15:48  Type: P  Check: 2629  ..

       1  MKCSWVMFFL MAVVTGVNSE VQLQQSGAEL VKPGASVKLS CTASGFNIKD

      51  TYIHCVKQRP EQGLEWIGRI DPANGYTKYD PKFQGKATIT ADTSSNTAYL

     101  QLSSLTSEDT AVYFCAREGY YGNYGVYAMD YWGQGTSVTV

!!AA_SEQUENCE 1.0
ID   R81330 standard; Protein; 123 AA.
AC   R81330;
DT   02-APR-1996  (first entry)
DE   Mouse anti-VLA-4 antibody 21.6 heavy chain variable region.
KW   Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
KW   antibody engineering.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   region          1. .30
FT                   /label= FR1
FT                   /note= "mouse heavy chain variable framework
FT                   region 1"
FT   region          31. .35
FT                   /label= CDR1
FT                   /note= "mouse heavy chain variable complementarity
FT                   determining region 1"
FT   region          36. .49
FT                   /label= FR2
FT                   /note= "mouse heavy chain variable framework
FT                   region 2"
FT   region          50. .66
FT                   /label= CDR2
```

```
FT                   /note= "mouse heavy chain variable complementarity
FT                   determining region 2"
FT   region          67. .98
FT                   /label= FR3
FT                   /note= "mouse heavy chain variable framework
FT                   region 3"
FT   region          99. .112
FT                   /label= CDR3
FT                   /note= "mouse heavy chain variable complementarity
FT                   determining region 3"
FT   region          113. .123
FT                   /label= FR4
FT                   /note= "mouse heavy light chain variable framework
FT                   region 4"
PN   WO9519790-A1.
PD   27-JUL-1995.
PF   25-JAN-1995; U01219.
PR   25-JAN-1994; US-186269.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones TS,  Leger OJ,  Saldanha J;
DR   WPI; 95-269276/35.
PT   New humanised antibodies against VLA-4 - used for inhibiting
PT   leukocyte adhesion to endothelial cells, partic. for treating
PT   inflammatory disease.
PS   Disclosure; Page 68; 105pp; English.
CC   The sequence represents the mouse anti-VLA-4 antibody 21.6 heavy chain
CC   variable region (without signal sequence). Cloned cDNA CDR sequences of
CC   mouse 21.6 variable light and variable heavy regions are linked to human
CC   constant framework regions of the REI antibody for the light chain and
CC   the 2*CL antibody for the heavy chain in the construction of a humanized
CC   antibody against VLA-4.  The 5' and 3' ends of the mouse cDNAs are
CC   modified using PCR primers (See Q99895-98) and then subcloned into
CC   mammalian cell expression vectors containing human kappa or gamma-1
CC   constant regions. In the humanized light chain, amino acids L45, L49,
CC   L58 and L69 in the human kappa LCVR framework are replaced by the amino
CC   acid present in the equivalent position of the mouse 21.6 Ig light chain.
CC   Plasmids encoding the chimeric antibodies are transfected into COS cells.
CC   The humanized antibodies can be used to inhibit adhesion of a leukocyte
CC   to an endothelial cell and to treat inflammatory diseases such as
CC   multiple sclerosis. They can also be used in the treatment of stroke,
CC   cerebral traumas, meningitis or encephalitis. The antibodies can also be
CC   used for detecting VLA-4, for affinity purification or for generating
CC   anti-idiotype antibodies.
SQ   Sequence   123 AA;

 R81330  Length: 123  February 11, 2000 15:48  Type: P  Check: 8013  ..

       1  EVQLQQSGAE LVKPGASVKL SCTASGFNIK DTYIHCVKQR PEQGLEWIGR

      51  IDPANGYTKY DPKFQGKATI TADTSSNTAY LQLSSLTSED TAVYFCAREG

     101  YYGNYGVYAM DYWGQGTSVT VSS

!!AA_SEQUENCE 1.0
ID   R81333 standard; Protein; 142 AA.
AC   R81333;
DT   23-MAR-1996  (first entry)
DE   Human VLA-4 reshaped antibody 21.6 light heavy variable region.
KW   Humanized antibody; leukocyte adhesion molecule; VLA-4; therapeutic;
KW   antibody engineering.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   peptide         1. .19
FT                   /note= "signal peptide"
FT   region          20. .49
FT                   /note= "framework region 1"
FT   region          50. .54
FT                   /note= "complementarity determining region 1"
FT   region          55. .68
FT                   /note= "framework region 2"
FT   region          69. .85
FT                   /note= "complementarity determining region 2"
FT   region          86. .117
FT                   /note= "framework region 3"
FT   region          118. .131
FT                   /note= "complementarity determining region 3"
FT   region          132. .142
FT                   /note= "framework region 4"
PN   WO9519790-A1.
PD   27-JUL-1995.
PF   25-JAN-1995; U01219.
PR   25-JAN-1994; US-186269.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones TS,  Leger OJ,  Saldanha J;
DR   WPI; 95-269276/35.
DR   N-PSDB; Q99894.
PT   New humanised antibodies against VLA-4 - used for inhibiting
PT   leukocyte adhesion to endothelial cells, partic. for treating
PT   inflammatory disease.
PS   Disclosure; Fig 11; 105pp; English.
CC   The sequence represents the human reshaped antibody 21.6 heavy
CC   chain variable region against leukocyte adhesion molecule VLA-4.
CC   Cloned cDNA sequences of mouse 21.6 VH (Q99892) and VL (Q99889)
CC   regions are linked to human constant regions in the construction
CC   of a humanized antibody against VLA-4.  The 5' and 3' ends of the
CC   mouse cDNAs are modified using PCR primers (Q99895-98) and then
CC   subcloned into mammalian cell expression vectors containing human
CC   kappa or gamma-1 constant regions.  In the humanized heavy chain,
CC   amino acids H27, H28, H29, H30, H44 nd H71 in the human HC VR
CC   framework are replaced by the amino acid present in the equivalent
CC   position of the mouse 21.6 Ig H chain.  Plasmids encoding the
CC   chimeric antibodies are transfected into COS cells.  The humanized
CC   antibodies can be used to inhibit adhesion of a leukocyte to an
CC   endothelial cell and to treat inflammatory diseases such as multiple
CC   sclerosis.  They can also be used in the treatment of stroke,
CC   cerebral traumas, meningitis or encephalitis.  The antibodies can
CC   also be used for detecting VLA-4, for affinity purification or for
CC   generating anti-idiotype antibodies.
SQ   Sequence   142 AA;

 R81333  Length: 142  February 11, 2000 15:48  Type: P  Check: 9019  ..

       1  MDWTWRVFCL LAVAPGAHSQ VQLVQSGAEV KKPGASVKVS CKASGFNIKD

      51  TYIHWVRQAP GQRLEWMGRI DPANGYTKYD PKFQGRVTIT ADTSASTAYM

     101  ELSSLRSEDT AVYYCAREGY YGNYGVYAMD YWGQGTLVTV SS

!!AA_SEQUENCE 1.0
ID   R75202 standard; Protein; 225 AA.
AC   R75202;
DT   10-MAY-1996  (first entry)
DE   Tyrosine phosphatase MPTP-delta insertion sequence, INS-1.
KW   Tyrosine phosphatase MPTP-delta; murine; brain tissue;
KW   glutathione-S-transferase; fusion protein; E. coli; differentiation;
KW   activation; information transmission; nervous system; immune system;
KW   carcinogenesis; insetion; INS-1; INS-2.
OS   Mus musculus domesticus.
PN   J07236487-A.
PD   12-SEP-1995.
PF   28-FEB-1994; 054726.
PR   28-FEB-1994; JP-054726.
PA   (TOKS-) TOKYOTO SHINKEI KAGAKU SOGO KENKYUSHO ZH.
DR   WPI; 95-347455/45.
DR   N-PSDB; Q94312.
PT   DNA encoding tyrosine phosphatase MPTP delta - useful for
PT   elucidation of signal transmission mechanisms.
PS   Claim 2; Page 11-12; 14pp; Japanese.
CC   The sequences given in R75202-03 are encoded by insertion sequences
CC   which were included in the tyrosine phosphatase MPTP-delta coding
CC   sequence.  The INS-1 cDNA was inserted between the 66th A and the 67th
CC   G of the MPTP-delta coding sequence, and the INS-2 cDNA was inserted
CC   between the 1194th A and the 1195th T of the sequence.  The MPTP-delta
CC   sequence was isolated from murine brain tissue and was cloned, for
CC   expression, into the downstream region of a glutathione-S-transferase
CC   sequence and expressed as a fusion protein in E. coli.  MPTP-delta
```

```
CC   proteins regulate differentiation and activation of cells.  This
CC   sequence can be used in the elucidation of the molecular mechanism for
CC   information transmission in cells, regulation mechanisms in the nervous
CC   system or immune system, or in the mechanism of carcinogenesis.
SQ   Sequence   225 AA;

 R75202  Length: 225  February 11, 2000 15:48  Type: P  Check: 824  ..

       1  MPLTNCRMVP VRRPLSLLLT FFLCACAETP PRFTRTPVDQ TGVSGGVASF

      51  ICQATGDPRP KIVWNKKGKK VSNQRFEVIE FDDGSGSVLR IQPLRTPRDE

     101  AIYECVASNN VGEISVSTRL TVLREDQIPR GFPTIDMGPQ LKVVERTRTA

     151  TMLCAASGNP DPEITWFKDF LPVDTSNNNG RIKQLRSESI GALQIEQSEE

     201  SDQGKYECVA TNSAGTRYSA PANLY

!!AA_SEQUENCE 1.0
ID   R91274 standard; Protein; 62 AA.
AC   R91274;
DT   18-JUN-1996  (first entry)
DE   Eimeria gametocyte antigen.
KW   Vaccine; coccidiosis; chicken; hen; fowl; immunisation;
KW   maternal immunity; gametocyte; antigen.
OS   Eimeria maxima.
PN   US5496550-A.
PD   05-MAR-1996.
PF   14-AUG-1986; 896611.
PR   14-AUG-1986; US-896611.
PR   12-FEB-1988; US-155245.
PR   14-FEB-1989; US-310603.
PR   16-JAN-1991; US-642219.
PR   17-AUG-1993; US-108763.
PA   (CHIL-) CHILWALNER.
PI   Mencher D,  Pugatsch T,  Wallach M;
DR   WPI; 96-150679/15.
DR   N-PSDB; T13968.
PT   Protection of newborn chicks from Eimeria spp. infection - by
PT   immunising a laying hen prior to the hen laying with an Eimeria spp.
PT   gametocyte antigenic protein
PS   Example 9; Fig 14a; 39pp; English.
CC   Clone pEM 56/3 (ATCC 68228) represents the 3' half of either a 2.0
CC   or 3.5 kb mRNA of Eimeria maxima gametocytes.  It was obtd. by
CC   screening an E. maxima cDNA library using mouse polyclonal antibodies
CC   to the 82 kDa gametocyte antigen.  There are 2 possible open reading
CC   frames (T13968 and T13969) coding for different proteins (R91274 and
CC   R91275, respectively).  Cloning of Eimeria gametocyte antigen genes
CC   (see also R91271-73) will allow large-scale prodn. of the antigens,
CC   useful for immunising hens to protect newborn chicks against
CC   coccidiosis.
SQ   Sequence   62 AA;

 R91274  Length: 62  February 11, 2000 15:48  Type: P  Check: 1456  ..

       1  LQHHHQHIIL QQQQHHHHPP PAAVLVAAPP ADKEQFAAAA AAAPIPIAAA

      51  SKETPLIDMM SK

!!AA_SEQUENCE 1.0
ID   R94989 standard; Peptide; 354 AA.
AC   R94989;
DT   06-JUL-1996  (first entry)
DE   Nsk2 receptor intracellular domain (aa518-871).
KW   Nsk2; neural fold/somite kinase; receptor tyrosine kinase; muscle;
KW   vector; tumour; muscular dystrophy; diagnosis; therapy.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   region          1. .59
FT                   /label= Juxtamembrane_region
FT   region          60. .341
FT                   /label= Tyrosine-kinase_region
FT   region          157. .177
FT                   /label= Kinase_insert
FT   region          342. .354
FT                   /label= C-terminal_region
PN   WO9611664-A2.
PD   25-APR-1996.
PF   10-OCT-1995; U13490.
PR   10-OCT-1994; GB-020389.
PR   24-MAY-1995; GB-010574.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Reith A;
DR   WPI; 96-221725/22.
PT   Nucleic acid encoding muscle receptor tyrosine kinase(s) - and
PT   related vectors, transformed cells, polypeptide(s) and antibodies,
PT   useful for treating tumours, diagnosing muscular dystrophy, etc.
PS   Claim 8; Fig 2; 65pp; English.
CC   The intracellular domain (R94989) and other polypeptide portions
CC   (see also R94981-88 and R94990-92) of mouse Nsk2 (R84087) can be
CC   used to study the role of Nsk2 in growth and differentiation, in
CC   the development of Nsk (ant)agonists, to identify other components
CC   of Nsk signalling pathways, and in the diagnosis and therapy of e.g.
CC   tumours and muscular dystrophy.  Nsk2 is a novel member of a new
CC   family of muscle receptor tyrosine kinases that are expressed in
CC   terminally differentiated skeletal myotubes in vitro and in adult
CC   skeletal muscles in vivo.
SQ   Sequence   354 AA;

 R94989  Length: 354  February 11, 2000 15:48  Type: P  Check: 5431  ..

       1  YCCRRRKEWK NKKRESTAVT LTTLPSELLL DRLHPNPMYQ RMPLLLNPKL

      51  LSLEYPRNNI EYVRDIGEGA FGRVFQARAP GLLPYEPFTM VAVKMLKEEA

     101  SADMQADFQR EAALMAEFDN PNIVKLLGVC AVGKPMCLLF EYMAYGDLNE

     151  FLRSMSPHTV CSLSHSDLST RARVSSPGPP PLSCAEQLCI ARQVAAGMAY

     201  LSERKFVHRD LATRNCLVGE TMVVKIADFG LSRNIYSADY YKADGNDAIP

     251  IRWMPPESIF YNRYTTESDV WAYGVVLWEI FSYGLQPYYG MAHEEVIYYV

     301  RDGNILACPE NCPLELYNLM RLCWSKLPAD RPSFCSIHRI LQRMCERAEG

     351  TVGV

!!AA_SEQUENCE 1.0
ID   R94990 standard; Peptide; 282 AA.
AC   R94990;
DT   06-JUL-1996  (first entry)
DE   Nsk2 receptor tyrosine-kinase region (aa577-858).
KW   Nsk2; neural fold/somite kinase; receptor tyrosine kinase; muscle;
KW   vector; tumour; muscular dystrophy; diagnosis; therapy.
OS   Mus musculus.
PN   WO9611664-A2.
PD   25-APR-1996.
PF   10-OCT-1995; U13490.
PR   10-OCT-1994; GB-020389.
PR   24-MAY-1995; GB-010574.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Reith A;
DR   WPI; 96-221725/22.
PT   Nucleic acid encoding muscle receptor tyrosine kinase(s) - and
PT   related vectors, transformed cells, polypeptide(s) and antibodies,
PT   useful for treating tumours, diagnosing muscular dystrophy, etc.
PS   Claim 8; Fig 2; 65pp; English.
CC   The tyrosine kinase region (R94990) and other polypeptide portions
CC   (see also R94981-89 and R94991-92) of mouse Nsk2 (R84087) can be
CC   used to study the role of Nsk2 in growth and differentiation, in
CC   the development of Nsk (ant)agonists, to identify other components
CC   of Nsk signalling pathways, and in the diagnosis and therapy of e.g.
CC   tumours and muscular dystrophy.  Nsk2 is a novel member of a new
CC   family of muscle receptor tyrosine kinases that are expressed in
CC   terminally differentiated skeletal myotubes in vitro and in adult
CC   skeletal muscles in vivo.
```

```
SQ   Sequence   282 AA;

 R94990  Length: 282  February 11, 2000 15:48  Type: P  Check: 5928  ..

       1  IEYVRDIGEG AFGRVFQARA PGLLPYEPFT MVAVKMLKEE ASADMQADFQ

      51  REAALMAEFD NPNIVKLLGV CAVGKPMCLL FEYMAYGDLN EFLRSMSPHT

     101  VCSLSHSDLS TRARVSSPGP PPLSCAEQLC IARQVAAGMA YLSERKFVHR

     151  DLATRNCLVG ETMVVKIADF GLSRNIYSAD YYKADGNDAI PIRWMPPESI

     201  FYNRYTTESD VWAYGVVLWE IFSYGLQPYY GMAHEEVIYY VRDGNILACP

     251  ENCPLELYNL MRLCWSKLPA DRPSFCSIHR IL

!!AA_SEQUENCE 1.0
ID   R89896 standard; Protein; 423 AA.
AC   R89896;
DT   14-JUL-1996  (first entry)
DE   Rat kynurenine aminotransferase (KAT) clone.
KW   Kynurenine aminotransferase; KAT; kynurenic acid; KYNA; kynurenine;
KW   KYN; brain; NMDA receptor; glutamatergic function.
OS   Rattus rattus.
PN   WO9601893-A1.
PD   25-JAN-1996.
PF   23-JUN-1995; U07855.
PR   07-JUL-1994; US-271667.
PA   (PHAA ) PHARMACIA SPA.
PA   (UYMA-) UNIV MARYLAND BALTIMORE.
PI   Benatti L,  Breton J,  Mosca M,  Okuno E,  Schwarcz R;
PI   Speciale C;
DR   WPI; 96-097623/10.
DR   N-PSDB; T11560.
PT   Isolated DNA encoding mammalian kynurenine amino:transferase (KAT) -
PT   useful in gene therapy applications and for identifying KAT in brain
PT   tissue
PS   Claim 16; Figure 2; 51pp; English.
CC   Sequences encoding Kynurenine aminotransferase (KAT) can be inserted
CC   into vectors and subsequently cells and hence can be used for gene
CC   therapy. The vector and host cells can be used for cerebral
CC   implantation to where KAT can directly catalyse the production of
CC   kynurenic acid (KYNA) from kynurenine (KYN). It is thought KYNA acts
CC   as a negative endogenous modulator of cerebral glutamatergic
CC   function. KYNA concentrations and the activity of KAT show an
CC   increase with age.  KAT inhibitors, by providing an increase of the
CC   glutamatergic tone at the NMDA receptor, could be useful in
CC   situations where NMDA receptor function is insufficient and/or KAT
CC   activity and KYNA levels are abnormally enhanced.  Hence they could
CC   be particularly useful in the treatment of the pathological
CC   consequences associated with the aging processes in the brain.
CC   Three KAT clones are described in T11560, T11742-43.
SQ   Sequence   423 AA;

 R89896  Length: 423  February 11, 2000 15:48  Type: P  Check: 6122  ..

       1  MTKRLQARRL DGIDQNLWVE FGKLTKEYDV VNLGQGFPDF SPPDFATQAF

      51  QQATSGNFML NQYTRAFGYP PLTNVLASFF GKLLGQEMDP LTNVLVTVGA

     101  YGALFTAFQA LVDEGDEVII MEPAFDCYEP MTMMAGGCPV FVTLKPSPAP

     151  KGKLGASNDW QLDPAELASK FTPRTKVLVL NTPNNPLGKV FSRMELELVA

     201  NLCQQHDVVC ISDEVYQWLV YDGHQHVSIA SLPGMWDRTL TIGSAGKSFS

     251  ATGWKVGWVM GPDNIMKHLR TVHQNSIFHC PTQAQAAVAQ CFEREQQHFG

     301  QPSSYFLQLP QAMELNRDHM IRSLQSVGLK LWISQGSYFL IADISDFKSK

     351  MPDLPGAEDE PYDRRFAKWM IKNMGLVGIP VSTFFSRPHQ KDFDHYIRFC

     401  FVKDKATLQA MDERLRKWKE LQP

!!AA_SEQUENCE 1.0
ID   R89897 standard; Protein; 437 AA.
AC   R89897;
DT   14-JUL-1996  (first entry)
DE   Rat kynurenine aminotransferase (KAT) clone.
KW   Kynurenine aminotransferase; KAT; kynurenic acid; KYNA; kynurenine;
KW   KYN; brain; NMDA receptor; glutamatergic function.
OS   Rattus rattus.
PN   WO9601893-A1.
PD   25-JAN-1996.
PF   23-JUN-1995; U07855.
PR   07-JUL-1994; US-271667.
PA   (PHAA ) PHARMACIA SPA.
PA   (UYMA-) UNIV MARYLAND BALTIMORE.
PI   Benatti L,  Breton J,  Mosca M,  Okuno E,  Schwarcz R;
PI   Speciale C;
DR   WPI; 96-097623/10.
DR   N-PSDB; T11742.
PT   Isolated DNA encoding mammalian kynurenine amino:transferase (KAT) -
PT   useful in gene therapy applications and for identifying KAT in brain
PT   tissue
PS   Claim 16; Figure 3; 51pp; English.
CC   Sequences encoding Kynurenine aminotransferase (KAT) can be inserted
CC   into vectors and subsequently cells and hence can be used for gene
CC   therapy. The vector and host cells can be used for cerebral
CC   implantation to where KAT can directly catalyse the production of
CC   kynurenic acid (KYNA) from kynurenine (KYN). It is thought KYNA acts
CC   as a negative endogenous modulator of cerebral glutamatergic
CC   function. KYNA concentrations and the activity of KAT show an
CC   increase with age.  KAT inhibitors, by providing an increase of the
CC   glutamatergic tone at the NMDA receptor, could be useful in
CC   situations where NMDA receptor function is insufficient and/or KAT
CC   activity and KYNA levels are abnormally enhanced.  Hence they could
CC   be particularly useful in the treatment of the pathological
CC   consequences associated with the aging processes in the brain.
CC   Three KAT clones are described in T11560, T11742-43.
SQ   Sequence   437 AA;

 R89897  Length: 437  February 11, 2000 15:48  Type: P  Check: 206  ..

       1  MNHSTWCLFF KKLTMTKRLQ ARRLDGIDQN LWVEFGKLTK EYDVVNLGQG

      51  FPDFSPPDFA TQAFQQATSG NFMLNQYTRA FGYPPLTNVL ASFFGKLLGQ

     101  EMDPLTNVLV TVGAYGALFT AFQALVDEGD EVIIMEPAFD CYEPMTMMAG

     151  GCPVFVTLKP SPAPKGKLGA SNDWQLDPAE LASKFTPRTK VLVLNTPNNP

     201  LGKVFSRMEL ELVANLCQQH DVVCISDEVY QWLVYDGHQH VSIASLPGMW

     251  DRTLTIGSAG KSFSATGWKV GWVMGPDNIM KHLRTVHQNS IFHCPTQAQA

     301  AVAQCFEREQ QHFGQPSSYF LQLPQAMELN RDHMIRSLQS VGLKLWISQG

     351  SYFLIADISD FKSKMPDLPG AEDEPYDRRF AKWMIKNMGL VGIPVSTFFS

     401  RPHQKDFDHY IRFCFVKDKA TLQAMDERLR KWKELQP

!!AA_SEQUENCE 1.0
ID   R89898 standard; Protein; 457 AA.
AC   R89898;
DT   14-JUL-1996  (first entry)
DE   Rat kynurenine aminotransferase (KAT) clone.
KW   Kynurenine aminotransferase; KAT; kynurenic acid; KYNA; kynurenine;
KW   KYN; brain; NMDA receptor; glutamatergic function.
OS   Rattus rattus.
PN   WO9601893-A1.
PD   25-JAN-1996.
PF   23-JUN-1995; U07855.
PR   07-JUL-1994; US-271667.
PA   (PHAA ) PHARMACIA SPA.
PA   (UYMA-) UNIV MARYLAND BALTIMORE.
```

```
PI   Benatti L,  Breton J,  Mosca M,  Okuno E,  Schwarcz R;
PI   Speciale C;
DR   WPI; 96-097623/10.
DR   N-PSDB; T11743.
PT   Isolated DNA encoding mammalian kynurenine amino:transferase (KAT) -
PT   useful in gene therapy applications and for identifying KAT in brain
PT   tissue
PS   Claim 16; Figure 4; 51pp; English.
CC   Sequences encoding Kynurenine aminotransferase (KAT) can be inserted
CC   into vectors and subsequently cells and hence can be used for gene
CC   therapy. The vector and host cells can be used for cerebral
CC   implantation to where KAT can directly catalyse the production of
CC   kynurenic acid (KYNA) from kynurenine (KYN). It is thought KYNA acts
CC   as a negative endogenous modulator of cerebral glutamatergic
CC   function. KYNA concentrations and the activity of KAT show an
CC   increase with age.  KAT inhibitors, by providing an increase of the
CC   glutamatergic tone at the NMDA receptor, could be useful in
CC   situations where NMDA receptor function is insufficient and/or KAT
CC   activity and KYNA levels are abnormally enhanced.  Hence they could
CC   be particularly useful in the treatment of the pathological
CC   consequences associated with the aging processes in the brain.
CC   Three KAT clones are described in T11560, T11742-43.
SQ   Sequence   457 AA;

 R89898  Length: 457  February 11, 2000 15:48  Type: P  Check: 7276  ..

       1  MFRSAAALSV HLMWPLWGRK AGASLTRCLH QSLTMTKRLQ ARRLDGIDQN

      51  LWVEFGKLTK EYDVVNLGQG FPDFSPPDFA TQAFQQATSG NFMLNQYTRA

     101  FGYPPLTNVL ASFFGKLLGQ EMDPLTNVLV TVGAYGALFT AFQALVDEGD

     151  EVIIMEPAFD CYEPMTMMAG GCPVFVTLKP SPAPKGKLGA SNDWQLDPAE

     201  LASKFTPRTK VLVLNTPNNP LGKVFSRMEL ELVANLCQQH DVVCISDEVY

     251  QWLVYDGHQH VSIASLPGMW DRTLTIGSAG KSFSATGWKV GWVMGPDNIM

     301  KHLRTVHQNS IFHCPTQAQA AVAQCFEREQ QHFGQPSSYF LQLPQAMELN

     351  RDHMIRSLQS VGLKLWISQG SYFLIADISD FKSKMPDLPG AEDEPYDRRF

     401  AKWMIKNMGL VGIPVSTFFS RPHQKDFDHY IRFCFVKDKA TLQAMDERLR

     451  KWKELQP

!!AA_SEQUENCE 1.0
ID   R77433 standard; Protein; 160 AA.
AC   R77433;
DT   19-AUG-1996  (first entry)
DE   Antigen of Rochalimaea henselae.
KW   Rochalimaea henselae; cat scratch disease; bacillary angiomatosis;
KW   CSD; infection; antigen; antibody; vaccine.
OS   Rochalimaea henselae.
PN   WO9531549-A1.
PD   23-NOV-1995.
PF   18-MAY-1995; U06211.
PR   18-MAY-1994; US-245294.
PR   16-SEP-1994; US-307279.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Anderson BE,  Regnery RL;
DR   WPI; 96-010935/01.
DR   N-PSDB; T04402.
PT   Nucleic acids of Rochalimaea henselae and R.quintana - methods which
PT   enable the identification of R.henselae, which is a causative agent
PT   of both cat scratch disease and bacillary angiomatosis
PS   Claim 7; Page 88; 100pp; English.
CC   The nucleic acids (T04402, T04403), fragments and antibodies
CC   binding to the encoded proteins (R77433, R77434), may be used in the
CC   diagnosis and detection of cat scratch disease (CSD) and bacillary
CC   angiomatosis caused by R. henselae.  The proteins or fragments of
CC   them, may be used in vaccines to protect against R. henselae
CC   infection.
SQ   Sequence   160 AA;

 R77433  Length: 160  February 11, 2000 15:48  Type: P  Check: 2940  ..

       1  MAAYISSKER KSMKKYSLVT LLSLFCISHA KAQTATLTDE YYKKALENTQ

      51  KLDVAKSQTA ESIYESATQT ANKIKDINNQ LANLKADTKT KPEQLQALQI

     101  ELTLLQAQLQ ADTLKIQSLA MIQAKDTKTK EELREEQTQK KHEDLQKQLK

     151  EKLEKSDVRL

!!AA_SEQUENCE 1.0
ID   R95054 standard; Protein; 342 AA.
AC   R95054;
DT   19-AUG-1996  (first entry)
DE   TGF-a-DETA-DGAL4 multidomain protein.
KW   Nucleic acid transfer system; gene transfer; gene therapy;
KW   cell targeting; multidomain protein; vector; cancer;
KW   exotoxin A; DETA; ompA; signal peptide; GAL4; TGF-a;
KW   transforming growth factor-alpha.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens;
OS   Chimeric Pseudomonas aeruginosa;
OS   Chimeric Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   peptide         1. .8
FT                   /label= FLAG_epitope
FT   peptide         9. .12
FT                   /label= Spacer
FT   domain          13. .62
FT                   /label= TGF-a
FT                   /note= "amino acids 1-50 of human TGF-a
FT   peptide         63. .65
FT                   /label= Spacer
FT   peptide         66. .71
FT                   /label= Hexa-histidine
FT   peptide         72
FT                   /label= Spacer
FT   domain          73. .187
FT                   /label= ETA
FT                   /note= "amino acids 252-366 of ETA"
FT   peptide         188
FT                   /label= Spacer
FT   domain          189. .334
FT                   /label= GAL4
FT                   /note= "amino acids 2-147 of yeast GAL4"
FT   peptide         335. .342
FT                   /label= Spacer
FT                   /note= "endoplasmic reticulum retention signal"
FT
PN   WO9613599-A1.
PD   09-MAY-1996.
PF   31-OCT-1995; E04270.
PR   01-NOV-1994; EP-810627.
PA   (WELS/) WELS W.
PI   Fominaya J,  Wels W;
DR   WPI; 96-239505/24.
DR   N-PSDB; T29409.
PT   Nucleic acid transfer system for gene therapy, e.g. against cancer
PT   - includes toxin translocation domain to target nucleic acid to
PT   specific cell
PS   Claim 7; Page 64-65; 106pp; English.
CC   A multidomain protein (R95054) has a FLAG epitope, a portion
CC   of human transforming growth factor-alpha (TGF-a) that acts as a
CC   ligand domain, a non-cytotoxic portion of Pseudomonas aeruginosa
CC   exotoxin A acting as a translocation domain and the DNA
CC   binding domain of yeast GAL4.  It is the product of a fusion
CC   gene (T29410) and can be expressed in E. coli (resulting in
CC   removal of an ompA signal peptide).  It is used with an effector
CC   nucleic acid that comprises e.g. a gene to be delivered to
CC   a cell and a cognate structure for the GAL4 DNA binding domain.
CC   This provides a novel means of nucleic acid transfer, suitable
CC   for gene therapy.
```

```
SQ   Sequence   342 AA;

 R95054  Length: 342  February 11, 2000 15:48  Type: P  Check: 9286  ..

       1  DYKDDDDKLG TGVVSHFNDC PDSHTQFCFH GTCRFLVQED KPACVCHSGY

      51  VGARCEHADL LASLEHHHHH HLEGGSLAAL TAHQACHLPL ETFTRHRQPR

     101  GWEQLEQCGY PVQRLVALYL AARLSWNQVD QVIRNALASP GSGGDLGEAI

     151  REQPEQARLA LTLAAAESER FVRQGTGNDE AGAANADEKL LSSIEQACDI

     201  CRLKKLKCSK EKPKCAKCLK NNWECRYSPK TKRSPLTRAH LTEVESRLER

     251  LEQLFLLIFP REDLDMILKM DSLQDIKALL TGLFVQDNVN KDAVTDRLAS

     301  VETDMPLTLR QHRISATSSS EESSNKGQRQ LTVSSSDYKD EL

!!AA_SEQUENCE 1.0
ID   R95055 standard; Protein; 421 AA.
AC   R95055;
DT   19-AUG-1996  (first entry)
DE   IL-2-DETA-DGAL4 multidomain protein.
KW   Nucleic acid transfer system; gene transfer; gene therapy;
KW   cell targeting; multidomain protein; vector; cancer;
KW   exotoxin A; DETA; ompA; signal peptide; GAL4; interleukin-2;
KW   IL-2.
OS   Chimeric synthetic;
OS   Chimeric Homo sapiens;
OS   Chimeric Pseudomonas aeruginosa;
OS   Chimeric Saccharomyces cerevisiae.
FH   Key             Location/Qualifiers
FT   peptide          1. .8
FT                    /label= FLAG_epitope
FT   peptide          9. .17
FT                    /label= Spacer
FT   domain           18. .150
FT                    /label= IL-2
FT                    /note= "amino acids 1-113 of human IL-2
FT   peptide          151
FT                    /label= Spacer
FT   domain           152. .266
FT                    /label= ETA
FT                    /note= "amino acids 252-366 of ETA"
FT   peptide          267
FT                    /label= Spacer
FT   domain           268. .413
FT                    /label= GAL4
FT                    /note= "amino acids 2-147 of yeast GAL4"
FT   peptide          414. .421
FT                    /label= Spacer
FT                    /note= "endoplasmic reticulum retention signal"
PN   WO9613599-A1.
PD   09-MAY-1996.
PF   31-OCT-1995; E04270.
PR   01-NOV-1994; EP-810627.
PA   (WELS/) WELS W.
PI   Fominaya J,  Wels W;
DR   WPI; 96-239505/24.
DR   N-PSDB; T29411.
PT   Nucleic acid transfer system for gene therapy, e.g. against cancer
PT   - includes toxin translocation domain to target nucleic acid to
PT   specific cell
PS   Claim 7; Page 67-69; 106pp; English.
CC   A multidomain protein (R95055) has a FLAG epitope, a portion
CC   of human interleukin-2 that acts as a ligand domain, a
CC   non-cytotoxic portion of Pseudomonas aeruginosa exotoxin A acting
CC   as a translocation domain and the DNA binding domain of yeast GAL4.
CC   It is the product of a fusion gene (T29411) and can be expressed
CC   in E. coli (resulting in removal of an ompA signal peptide).  It is
CC   used with an effector nucleic acid that comprises e.g. a gene to be
CC   delivered to a cell and a cognate structure for the GAL4 DNA binding
CC   domain.  This provides a novel means of nucleic acid transfer,
CC   suitable for gene therapy.
SQ   Sequence   421 AA;

 R95055  Length: 421  February 11, 2000 15:48  Type: P  Check: 7433  ..

       1  DYKDDDDKLH HHHHHKLAPT SSSTKKTQLQ LEHLLLDLQM ILNGINNYKN

      51  PKLTRMLTFK FYMPKKATEL KHLQCLEEEL KPLEEVLNLA QSKNFHLRPR

     101  DLISNINVIV LELKGSETTF MCEYADETAT IVEFLNRWIT FCQSIISTLT

     151  LEGGSLAALT AHQACHLPLE TFTRHRQPRG WEQLEQCGYP VQRLVALYLA

     201  ARLSWNQVDQ VIRNALASPG SGGDLGEAIR EQPEQARLAL TLAAAESERF

     251  VRQGTGNDEA GAANADEKLL SSIEQACDIC RLKKLKCSKE KPKCAKCLKN

     301  NWECRYSPKT KRSPLTRAHL TEVESRLERL EQLFLLIFPR EDLDMILKMD

     351  SLQDIKALLT GLFVQDNVNK DAVTDRLASV ETDMPLTLRQ HRISATSSSE

     401  ESSNKGQRQL TVSSSDYKDE L

!!AA_SEQUENCE 1.0
ID   R91225 standard; Protein; 328 AA.
AC   R91225;
DT   26-AUG-1996  (first entry)
DE   Human placenta G-protein coupled receptor protein.
KW   G-protein coupled receptor protein; G-PCRP; agonist; antagonist;
KW   cystic fibrosis; incontinence; diabetes; diagnosis; therapy.
OS   Homo sapiens.
PN   WO9605302-A1.
PD   22-FEB-1996.
PF   10-AUG-1995; J01599.
PR   11-AUG-1994; JP-189272.
PR   11-AUG-1994; JP-189273.
PR   11-AUG-1994; JP-189274.
PR   30-SEP-1994; JP-236356.
PR   30-SEP-1994; JP-236357.
PR   02-NOV-1994; JP-270017.
PR   28-DEC-1994; JP-326611.
PR   20-JAN-1995; JP-007177.
PR   16-MAR-1995; JP-057186.
PR   19-APR-1995; JP-093989.
PA   (TAKE ) TAKEDA CHEM IND LTD.
PI   Fujii R,  Fukusumi S,  Hinuma S,  Hosoya M,  Ohgi K;
PI   Ohtaki T;
DR   WPI; 96-139698/14.
DR   N-PSDB; T18368.
PT   G-protein coupled receptor protein DNA and protein - also methods
PT   for isolating (ant)agonists for treatment of cystic fibrosis,
PT   incontinence and diabetes
PS   Claim 6; Page 267-68; 360pp; English.
CC   A novel human placenta-derived G-protein coupled receptor protein
CC   (G-PCRP) (R91225) was identified as the product of cDNA clone
CC   phAH2-17 (T18368).  The protein can be obtd. by expression of the
CC   cDNA clone in transformed host cells.  It was classified as a
CC   purinoceptor.  G-PCRPs (see also R91217-24 and R91227-33) can be
CC   used to screen agonists and antagonists that modulate G-PCRP
CC   activity, to raise antibodies and to develop assay systems.
SQ   Sequence   328 AA;

 R91225  Length: 328  February 11, 2000 15:48  Type: P  Check: 7811  ..

       1  MEWDNGTGQA LGLPPTTCVY RENFKQLLLP PVYSAVLAAG LPLNICVITQ

      51  ICTSRRALTR TAVYTLNLAL ADLLYACSLP LLIYNYAQGD HWPFGDFACR

     101  LVRFLFYANL HGSILFLTCI SFQRYLGICH PLAPWHKRGG RRAAWLVCVT

     151  VWLAVTTQCL PTAIFAATGI QRNRTVCYDL SPPALATHYM PYGMALTVIG

     201  FLLPFAALLA CYCLLACRLC RQDGPAEPVA QERRGKAARM AVVVAAAFAI
```

251 SFLPFHITKT AYLAVGSTPG VPCTVLEAFA AAYKGTRPFA SANSVLDPIL

301 FYFTQKKFRR RPHELLQKLT AKWQRQGR

!!AA_SEQUENCE 1.0
ID R99695 standard; Peptide; 25 AA.
AC R99695;
DT 25-SEP-1996 (first entry)
DE VP6 binding peptide B.
KW VP6; binding peptide; encapsulation; target; agent; drug delivery;
KW virus protein 6; rotavirus; specific.
OS Synthetic.
PN PT-100090-A.
PD 30-JUL-1993.
PF 23-DEC-1992; 100090.
PR 04-FEB-1991; US-650054.
PR 29-JAN-1992; US-825522.
PR 06-SEP-1994; US-301267.
PA (UYSA-) UNIV SASKATCHEWAN.
PI Redmond MJ, Santos M, Campos M;
DR WPI; 93-270390/34.
PT Encapsulated bioactive protein compsns. - contg. protein in
PT rotavirus VP6 protein spheres
PS Disclosure; Column 9; 12pp; English.
CC VP6 (is a 45 kD protein found in the inner capsid of rotaviruses. VP6
CC rapidly targets, attaches to and is engulfed by monocytes and
CC macrophages. Substances to be delivered to target cells can be
CC encapsulated in VP6 spheres to effect an increase in the concentration
CC of the agent delivered to the cell. Additionally, other targeting
CC agents can be attached to the VP6 spheres that will direct the spheres
CC to other tissues and cell types so that biologically active agents
CC encapsulated within the spheres can act locally. An advantage of the VP6
CC protein is that this system facilitates the attachment of molecules with
CC minimal manipulation, through protein-protein interactions. A synthetic
CC peptide corresponding to the targeting agent of interest can be
CC chemically synthesised in such a way that it also contains an amino acid
CC sequence (binding peptide) necessary to link it to VP6. Attachment of the
CC molecule to the VP6 carrier is then simply achieved by mixing the 2
CC substances without additional manipulation. The present sequence is a
CC peptide that binds to VP6. It occurs naturally as a portion of VP4 of
CC rotaviruses and is sensitive to trypsin. Cleavage of the peptide by
CC trypsin prevents it from binding to VP6. Patent No. US5503833 was used
CC to compile this file.
SQ Sequence 25 AA;

R99695 Length: 25 February 11, 2000 15:48 Type: P Check: 4995 ..

1 CNIAPASIVS RNIVYTRAQP NQDIA

!!AA_SEQUENCE 1.0
ID R99501 standard; Protein; 244 AA.
AC R99501;
DT 14-SEP-1996 (first entry)
DE Nitrilase regulatory factor 244-amino acid subunit.
KW Nitrilase regulatory factor; gene activator; promoter; nitrile.
OS Rhodococcus erythropolis strain SK92 (FERM BP-3324).
PN EP-719862-A2.
PD 03-JUL-1996.
PF 27-DEC-1995; 309454.
PR 28-DEC-1994; JP-337652.
PA (NITT) NITTO CHEM IND CO LTD.
PI Mizumura Y, Yu F;
DR WPI; 96-302346/31.
DR N-PSDB; T32242.
PT Rhodococcus erythropolis two component nitrilase gene activator -
PT enhances activation of nitrilase gene promoter in presence of
PT nitrile
PS Claim 1; Page 9-10; 23pp; English.
CC The nitrilase regulatory factor of Rhodococcus erythropolis SK92
CC has 2 subunits, of 244 amino acids (R99501) and 534 amino acids
CC (R95502). These are the products of the 2 open reading frames
CC (see also T32242 and T32243) of a gene (T32241) located upstream
CC of the nitrilase gene in SK92 chromosomal DNA. Introduction of a
CC plasmid carrying the regulatory gene plus the nitrilase gene and its
CC promoter allows efficient prodn. of nitrilase by a Rhodococcus sp.
CC transformant. The nitrilase is useful for the commercial prodn. of
CC organic acids and amides from nitriles.
SQ Sequence 244 AA;

R99501 Length: 244 February 11, 2000 15:48 Type: P Check: 2030 ..

1 MAGADVHAQG GTNRRARILV VDDEKHVRTM VTWQLESENF DVVAAADGDA

51 ALRQVTESGP DLMVLDLSLP GKGGLEVLAT VRRTDALPIV VLTARRDETE

101 RIVALDLGAD DYVIKPFSPR ELAARIRAVL RRTTAEPPHE AAVQRFGDLE

151 IDTAAREVRL HGIPLEFTTK EFDLLAYMAA SPMQVFSRRR LLLEVWRSSP

201 DWQQDATVTE HVHRIRRKIE EDPTKPTILQ TVRGAGYRFD GERA

!!AA_SEQUENCE 1.0
ID R99576 standard; Protein; 83 AA.
AC R99576;
DT 30-SEP-1996 (first entry)
DE Wasp venom BrhTX-1 subunit (b) partial sequence.
KW Wasp; venom; neurotoxin; insecticide; biological control agent;
KW Lepidoptera; insect.
OS Bracon hebetor.
PN WO9616171-A1.
PD 30-MAY-1996.
PF 21-NOV-1995; G02720.
PR 22-NOV-1994; GB-023540.
PR 19-JAN-1995; GB-001074.
PR 29-JUN-1995; GB-013293.
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (ZENE) ZENECA LTD.
PI Baule VJ, Christian PD, Duncan RE, Windass JD;
DR WPI; 96-268607/27.
DR N-PSDB; T32444.
PT Bracon hebetor toxins and DNA encoding them - useful in biological
PT control agents to combat insect pests
PS Example 6; Fig 11; 83pp; English.
CC Partial cDNA clone pBrhTX-1(b)1 (T32444) codes for a portion
CC (R99576) of subunit (b) of the Bracon hebetor wasp neurotoxin
CC BrhTX-1. It was obtd. by PCR amplification of wasp cDNA in
CC a phage lambda gt11 library and characterisation of a plaque-purified
CC phage clone. Primers BH(b) and BH(b)D (T32445-46) based on the clone
CC were used to re-screen the library, leading to the isolation of clone
CC BrhTX-1(b) (T32429) that coded for the toxin (b) subunit (R99577).
CC This can be utilised in breeding of biological control agents used to
CC combat insect pests.
SQ Sequence 83 AA;

R99576 Length: 83 February 11, 2000 15:48 Type: P Check: 2766 ..

1 FFLITMSIIC KIILLVLLSW TSMVSSTLFT DRKWCGRADK TFGPSRSLGG

51 GVGDCCRSHD SCGRMIKPGE TYGDVTNKGF SNM

!!AA_SEQUENCE 1.0
ID W00624 standard; Protein; 358 AA.
AC W00624;
DT 23-OCT-1996 (first entry)
DE SacI endonuclease.
KW SaclM; saclR; Streptomyces achromogenes; SacI; methylase; endonuclease.
OS Streptomyces achromogenes.
PN US5532153-A.
PD 02-JUL-1996.
PF 23-MAR-1995; 409199.
PR 23-MAR-1995; US-409199.
PA (NEWE) NEW ENGLAND BIOLABS INC.
PI Xiao J, Xu S;
DR WPI; 96-321120/32.
DR N-PSDB; T35805.

```
PT   DNA encoding SacI restriction endonuclease and methylase - useful
PT   for prodn. of the enzyme.
PS   Claim 3; Column 15-18; 16pp; English.
CC   This sequence represents the SacI endonuclease from Streptomyces
CC   achromogenes.  W00623 represents the SacI methylase.  The DNA encoding
CC   these sequences were isolated from isolated from genomic DNA of
CC   S.achromogenes, and amplified using the primers shown in T35806-T35811.
CC   The amplified sequences were then cloned into the plasmid pLG339 which
CC   was used to modify an E.coli host cell.  The host cell can then be used
CC   to produce SacI.
SQ   Sequence   358 AA;

 W00624  Length: 358  February 11, 2000 15:48  Type: P  Check: 8391  ..

       1  MGITIKKSTA EQVLRKAYEA AASDDVFLED WIFLATSLRE VDAPRTYTAA

      51  LVTALLARAC DDRVDPRSIK EKYDDRAFSL RTLCHGVVVP MSVELGFDLG

     101  ATGREPINNQ PFFRYDQYSE IVRVQTKARP YLDRVSSALA RVDEEDYSTE

     151  ESFRALVAVL AVCISVANKK QRVAVGSAIV EASLIAETQS FVVSGHDVPR

     201  KLQACVAAGL DMVYSEVVSR RINDPSRDFP GDVQVILDGD PLLTVEVRGK

     251  SVSWEGLEQF VSSATYAGFR RVALMVDAAS HVSLMSADDL TSALERKYEC

     301  IVKVNESVSS FLRDVFVWSP RDVHSILSAF PEAMYRRMIE IEVREPELDR

     351  WAEIFPET

!!AA_SEQUENCE 1.0
ID   R94579 standard; Protein; 488 AA.
AC   R94579;
DT   07-NOV-1996  (first entry)
DE   Chlamydia pneumoniae polypeptide antigen (polypeptide A).
KW   Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW   pCPN533alpha; primer; assay; detection; antibody; diagnosis;
KW   infection.
OS   Chlamydia pneumoniae.
FH   Key             Location/Qualifiers
FT   peptide         1. .259
FT                   /note= "claimed N-terminal fragment"
PN   WO9609320-A1.
PD   28-MAR-1996.
PF   20-SEP-1995; J01896.
PR   20-SEP-1994; JP-224711.
PR   28-APR-1995; JP-106011.
PR   28-APR-1995; JP-106006.
PR   28-APR-1995; JP-106010.
PR   28-APR-1995; JP-106008.
PR   28-APR-1995; JP-106009.
PA   (HITB ) HITACHI CHEM CO LTD.
PI   Izutsu H,  Matsumoto A,  Obara K;
DR   WPI; 96-188399/19.
DR   N-PSDB; T14612.
PT   Recombinant Chlamydia pneumoniae antigen and antibodies to it  -
PT   used for detection and assay of C. pneumoniae e.g. in clinical
PT   diagnosis
PS   Claim 1; Pages 60-64; 128pp; Japanese.
CC   The present sequence is the C. pneumoniae polypeptide antigen,
CC   polypeptide A. C. pneumoniae strain YK41 was cultured and genomic
CC   DNA extracted to prep. a lambda gt11 DNA library. The library was
CC   then screened with an anti-YK41 monoclonal antibody (MAb), which
CC   was prepd. by fusing spleen cells from a mouse infected with YK41
CC   with myeloma P3/NS1/1-Ag4-1 to produce a MAb expressing hydridoma.
CC   The DNA obtd. was then fused with the expression vector pADA431 to
CC   give pCPN533alpha. The plasmid was used to transform an E. coli
CC   host, which was cultured to give the antigenic polypeptide,
CC   polypeptide A. Polypeptide A and primers and probes derived from
CC   its DNA can be used in assays for the detection of polypeptide A
CC   antibodies and DNA, respectively, useful in the diagnosis of
CC   C. pneumoniae infection.
SQ   Sequence   488 AA;

 R94579  Length: 488  February 11, 2000 15:48  Type: P  Check: 2400  ..

       1  MSISSSSGPD NQKNIMSQVL TSTPQGVPQQ DKLSGNETKQ IQQTRQGKNT

      51  EMESDATIAG ASGKDKTSST TKTETAPQQG VAAGKESSES QKAGADTGVS

     101  GAAATTASNT ATKIAMQTSI EEASKSMEST LESLQSLSAA QMKEVEAVVV

     151  AALSGKSSGS AKLETPELPK PGVTPRSEVI EIGLALAKAI QTLGEATKSA

     201  LSNYASTQAQ ADQTNKLGLE KQAIKIDKER EEYQEMKAAE QKSKDLEGTM

     251  DTVNTVMIAV SVAITVISIV AAIFTCGAGL AGLAARAAVR AAAAGGAAGA

     301  AAATTVATQI TVQAVVQAVK QAVITAVRQA ITAAIKAAVK SGIKAFIKTL

     351  VKAIAKAISK GISKVFAKGT QMIAKNFPKL SKVISSLTSK WVTVGVGVVV

     401  AAPALGKGIM QMQLSEMQQN VAQFQKEVGK LQAAADMISM FTQFWQQASK

     451  IASKQTGESN EMTQKATKLG AQILKAYAAI SGAIAGAA

!!AA_SEQUENCE 1.0
ID   R94580 standard; Protein; 271 AA.
AC   R94580;
DT   07-NOV-1996  (first entry)
DE   C. pneumoniae polypeptide antigen (polypeptide A) variant.
KW   Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW   pCPN533alpha; primer; assay; detection; antibody; diagnosis;
KW   infection; variant.
OS   Chlamydia pneumoniae.
PN   WO9609320-A1.
PD   28-MAR-1996.
PF   20-SEP-1995; J01896.
PR   20-SEP-1994; JP-224711.
PR   28-APR-1995; JP-106011.
PR   28-APR-1995; JP-106006.
PR   28-APR-1995; JP-106010.
PR   28-APR-1995; JP-106008.
PR   28-APR-1995; JP-106009.
PA   (HITB ) HITACHI CHEM CO LTD.
PI   Izutsu H,  Matsumoto A,  Obara K;
DR   WPI; 96-188399/19.
DR   N-PSDB; T14613.
PT   Recombinant Chlamydia pneumoniae antigen and antibodies to it  -
PT   used for detection and assay of C. pneumoniae e.g. in clinical
PT   diagnosis
PS   Claim 2; Pages 64-66; 128pp; Japanese.
CC   The present sequence is a variant of the C. pneumoniae
CC   polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was
CC   cultured and genomic DNA extracted to prep. a lambda gt11 DNA
CC   library. The library was then screened with an anti-YK41
CC   monoclonal antibody (MAb), which was prepd. by fusing spleen cells
CC   from a mouse infected with YK41 with myeloma P3/NS1/1-Ag4-1 to
CC   produce a MAb expressing hydridoma. The DNA obtd. was then fused
CC   with the expression vector pADA431 to give pCPN533alpha. The
CC   plasmid was used to transform an E. coli host, which was cultured
CC   to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC   and primers and probes derived from its DNA can be used in assays
CC   for the detection of polypeptide A antibodies and DNA,
CC   respectively, useful in the diagnosis of C. pneumoniae infection.
SQ   Sequence   271 AA;

 R94580  Length: 271  February 11, 2000 15:48  Type: P  Check: 3773  ..

       1  MSISSSSGPD NQKNIMSQVL TSTPQGVPQQ DKLSGNETKQ IQQTRQGKNT

      51  EMESDATIAG ASGKDKTSST TKTETAPQQG VAAGKESSES QKAGADTGVS

     101  GAAATTASNT ATKIAMQTSI EEASKSMEST LESLQSLSAA QMKEVEAVVV

     151  AALSGKSSGS AKLETPELPK PGVTPRSEVI EIGLALAKAI QTLGEATKSA
```

201 LSNYASTQAQ ADQTNKLGLE KQAIKIDKER EEYQEMKAAE QKSKDLEGTM

251 DTVNTVMIAK GFELPWGPLI N

!!AA_SEQUENCE 1.0
ID R94586 standard; Protein; 259 AA.
AC R94586;
DT 11-NOV-1996 (first entry)
DE C. pneumoniae polypeptide antigen (polypeptide A) clone 53-3S.
KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW pCPN533alpha; primer; assay; detection; antibody; diagnosis;
KW infection; clone.
OS Chlamydia pneumoniae.
PN WO9609320-A1.
PD 28-MAR-1996.
PF 20-SEP-1995; J01896.
PR 20-SEP-1994; JP-224711.
PR 28-APR-1995; JP-106011.
PR 28-APR-1995; JP-106006.
PR 28-APR-1995; JP-106010.
PR 28-APR-1995; JP-106008.
PR 28-APR-1995; JP-106009.
PA (HITB) HITACHI CHEM CO LTD.
PI Izutsu H, Matsumoto A, Obara K;
DR WPI; 96-188399/19.
DR N-PSDB; T14622.
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it -
PT used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
PS Example 1; Pages 75-77; 128pp; Japanese.
CC The present sequence is the C. pneumoniae polypeptide antigen
CC polypeptide A clone, 53-3S. C. pneumoniae strain YK41 was
CC cultured and genomic DNA extracted to prep. a lambda gt11 DNA
CC library. The library was then screened with an anti-YK41
CC monoclonal antibody (MAb), which was prepd. by fusing spleen cells
CC from a mouse infected with YK41 with myeloma P3/NS1/1-Ag4-1 to
CC produce a MAb expressing hydridoma. The DNA obtd. was then fused
CC with the expression vector pADA431 to give pCPN533alpha. The
CC plasmid was used to transform an E. coli host, which was cultured
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC and primers and probes derived from its DNA can be used in assays
CC for the detection of polypeptide A antibodies and DNA,
CC respectively, useful in the diagnosis of C. pneumoniae infection.
SQ Sequence 259 AA;

R94586 Length: 259 February 11, 2000 15:48 Type: P Check: 9728 ..

1 MSISSSSGPD NQKNIMSQVL TSTPQGVPQQ DKLSGNETKQ IQQTRQGKNT

51 EMESDATIAG ASGKDKTSST TKTETAPQQG VAAGKESSES QKAGADTGVS

101 GAAATTASNT ATKIAMQTSI EEASKSMEST LESLQSLSAA QMKEVEAVVV

151 AALSGKSSGS AKLETPELPK PGVTPRSEVI EIGLALAKAI QTLGEATKSA

201 LSNYASTQAQ ADQTNKLGLE KQAIKIDKER EEYQEMKAAE QKSKDLEGTM

251 DTVNTVMIA

!!AA_SEQUENCE 1.0
ID R99316 standard; Protein; 458 AA.
AC R99316;
DT 13-NOV-1996 (first entry)
DE Human SH-PTP1 variant derived from erythroleukaemia cell line.
KW PTP; protein tyrosine phosphatase; SH2; Src homology region 2;
KW chromosome 12p; abnormality; mutation; detection; probe; neoplasia;
KW cancer; leukaemia; diagnosis; megakaryocyte regulation.
OS Homo sapiens.
PN US5536636-A.
PD 16-JUL-1996.
PF 26-JUN-1991; 721112.
PR 26-JUN-1991; US-721112.
PR 31-JAN-1992; US-829141.
PR 01-DEC-1992; US-983926.
PR 28-FEB-1994; US-202389.
PA (BETH-) BETH ISRAEL HOSPITAL ASSOC.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
PI Freeman RM, Neel BG, Plutzky J, Rosenberg RD;
DR WPI; 96-341506/34.
DR N-PSDB; T35310.
PT Detecting 12p chromosomal abnormality associated with neoplastic
PT disease - using SH-PTP1 protein tyrosine phosphatase gene specific
PT probe
PS Example 1; Column 49-52; 63pp; English.
CC R99316 is a variant SH-PTP1 (protein tyrosine phosphatase-1, with
CC two SH2 domains) obtained from a human erythroleukaemia cell line.
CC The variant cDNA sequence (see T35310) lacks a sequence of two
CC nucleotides, AG, corresponding to positions 1868 and 1869 of the
CC wild-type sequence (see T35306). A fragment complemetary to the
CC sequence from nucleotides 537-653 is used as a probe to detect a
CC chromosome 12p13 abnormality associated with neoplastic disease, in
CC partic. acute lymphoblastic leukaemia. The probe hybridises to a part
CC of the region coding for the two tandem SH2 domains (see R99312).
CC If the probe will not hybridise DNA of chromosome 12p13 from a patient
CC sample it is indicative of an abnormality, normally associated with
CC neoplasia. Alternatively the wild-type SH-PTP1 or SH-PTP2 gene or
CC protein may be used for comparison to sequenced PTP genes taken from a
CC patient, where differences indicate an abnormality. The activity of
CC SH-PTP1 may also be purposely altered by mutation to effect a change
CC in megakaryocyte function and hence platelet production.
SQ Sequence 458 AA;

R99316 Length: 458 February 11, 2000 15:48 Type: P Check: 4323 ..

1 MVRWFHRDLS GLDAETLLKG RGVHGSFLAR PSRKNQGDFS LSVRVGDQVT

51 HIRIQNSGDF YDLYGGEKFA TLTELVEYYT QQQGVVQDRD GTIIHLKYPL

101 NCSDPTSERW YHGHMSGGQA ETLLQAKGEP WTFLVRESLS QPGDFVLSVL

151 SDQPKAGPGS PLRVTHIKVM CEGGRYTVGG LETFDSLTDL VEHFKKTGIE

201 EASGAFVYLR QPYYATRVNA ADIENRVLEL NKKQESEDTA KAGFWEEFES

251 LQKQEVKNLH QRLEGQRPEN KGKNRYKNIL PFDHSRVILQ GRDSNIPGSD

301 YINANYIKNQ LLGPDENAKT YIASQGCLEA TVNDFWQMAW QENSRVIVMT

351 TREVEKGRNK CEYGNITYPP AMKNAHAKAS RTSSKHKEDV YENLHTKNKR

401 EESEEAAVSR QGEEQGFPQE EVSGAVLRWP CLSPDPVEAF RDGQTHNLNL

451 GVPHSFVI

!!AA_SEQUENCE 1.0
ID R83015 standard; Protein; 371 AA.
AC R83015;
DT 27-NOV-1996 (first entry)
DE Human thyroid transcription factor-1.
KW Thyroid transcription factor; TTF-1; human adenocarcinoma cell line;
KW H441; rat; mouse; pulmonary adenocarcinoma; H820; small cell carcinoma;
KW H345; tracheal-bronchial epithelial cell lines; respiratory epithelium;
KW fetal lung; gestation; pro-SP-C; respiratory epithelial cell;
KW nonciliated bronchiolar cell; immature lung; alveolar; Type II;
KW epithelial cell; nonciliated; Type I; adult.
OS Homo sapiens.
PN WO9531729-A1.
PD 23-NOV-1995.
PF 17-MAY-1995; U06244.
PR 18-MAY-1994; US-245356.
PA (CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.
PI Bohinski RJ, Whitsett JA;
DR WPI; 96-011078/01.
DR N-PSDB; T05810.
PT Oligo- and polynucleotide(s) that bind to lung cell nuclear proteins

PT - useful for cancer diagnosis and therapy.
PS Claim 50; Fig 39; 157pp; English.
CC This sequence represents the thyroid transcription factor, TTF-1. The
CC TTF-1 DNA sequence was isolated from the human adenocarcinoma cell line
CC H441. The TTF-1 locus is contained within a 4.6 kb BamHI fragment and
CC consists of two exons and one intron. This predicted amino acid
CC sequence of human TTF-1 shows close identity with the amino acid
CC sequence predicted by the rat TTF-1 cDNA sequence and 92.4% identity
CC with the rat TTF-1 cDNA. The intron is approx. 1 kb in length and
CC is flanked by consensus splice donor-acceptor sites that fit splice
CC acceptor-donor rules. The mRNA produced is 2.3 kb as detected by
CC Northern blot analysis of mRNA derived from mouse and human
CC adenocarcinoma cells. TTF-1 mRNA was detected in human pulmonary
CC adenocarcinoma cells H441 and H820 and small cell carcinoma H345, but
CC not detected in 9/HTEo- or BEAS-2B tracheal-bronchial epithelial cell
CC lines, A549, HeLA or 3T3 cells, demonstrating the cell selectivity of
CC TTF-1 expression. TTF-1 has been detected in nuclei of the respiratory
CC epithelium in human fetal lungs as early as 11-12 weeks of gestation.
CC Immunostaining has observed a distribution pattern in the developing
CC airways similar to the for pro-SP-C. TTF-1 was detected in subsets of
CC respiratory epithelial cells in the developing lung, including
CC nonciliated bronchiolar, and rarely in nonciliated bronchila respiratory
CC epithelial cells in the immature lung. At the time of birth TTF-1 was
CC detected in alveolar Type II epithelial cells and in subsets of
CC nonciliated bronchiolar epithelial cells. TTF-1 was not detected in
CC alveolar Type I cells or ciliated epithelial cells. In the adult lung,
CC TTF-1 is detected in subsets of nonciliated bronchiolar epithelial cells
CC and was most prominent in type II epithelial cells but was excluded from
CC Type I cells.
SQ Sequence 371 AA;

R83015 Length: 371 February 11, 2000 15:48 Type: P Check: 5439 ..

1 MSMSPKHTTP FSVSDILSPL EESYKKVGME GGGLGAPLAA YRQGQAAPPT

51 AAMQQHAVGH HGAVTAAYHM TAAGVPQLSH SAVGGYCNGN LGNMSELPPY

101 QDTMRNSASG PGWYGANPDP RFPAISRFMG PASGMNMSGM GGLGSLGDVS

151 KNMAPLPSAP RRKRRVLFSQ AQVYELERRF KQQKYLSAPE REHLASMIHL

201 TPTQVKIWFQ NHRYKMKRQA KDKAAQQQLQ QDSGGGGGGG GTGCPQQQQA

251 QQQSPRRVAV PVLVKDGKPC QAGAPAPGAA SLQGHAQQQA QHQAQAAQAA

301 AAAISVGSGG AGLGAHPGHQ PGSAGQSPDL AHHAASPAAL QGQVSSLSHL

351 NSSGSDYGTM SCSTLLYGRT W

!!AA_SEQUENCE 1.0
ID W00725 standard; Protein; 207 AA.
AC W00725;
DT 30-NOV-1996 (first entry)
DE Vascular endothelial growth factor-like protein SOM175.
KW Vascular endothelial growth factor; VEGF; VEGF165; SOM175; neuron;
KW astroglial proliferation.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1. .21
FT /label= Sig_peptide
PN WO9627007-A1.
PD 06-SEP-1996.
PF 22-FEB-1996; AU0094.
PR 02-MAR-1995; AU-001457.
PR 20-NOV-1995; AU-006647.
PR 22-DEC-1995; AU-007274.
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
PI Grimmond S, Hayward NK, Larsson C, Nordenskjold M;
PI Weber G;
DR WPI; 96-412774/41.
DR N-PSDB; T33610.
PT New growth factor related to vascular endothelial growth factor -
PT useful for inducing astroglial proliferation and promoting neuronal
PT survival
PS Claim 11; Page 41; 113pp; English.
CC Human vascular endothelial growth factor (VEGF)-like polypeptide
CC (W00725) is capable inducing the proliferation of vascular
CC endothelial cells, of interacting with flt-1/fkl-1 receptors,
CC and of inducing cell migration, cell survival and/or an increase
CC in intracellular levels of alkaline phosphatase. It shows 33.3%
CC identity with human VEGF (see also W00724). Splice variants
CC (W00726-28) of SOM175 have also been identified. Recombinant SOM175
CC can be produced in host cells transformed with vectors carrying
CC SOM175 cDNA (see also T33610). It is useful for inducing astroglial
CC proliferation and for promoting neural survival and/or proliferation.
SQ Sequence 207 AA;

W00725 Length: 207 February 11, 2000 15:48 Type: P Check: 1679 ..

1 MSPLLRRLLL AALLQLAPAQ APVSQPDAPG HQRKVVSWID VYTRATCQPR

51 EVVVPLTVEL MGTVAKQLVP SCVTVQRCGG CCPDDGLECV PTGQHQVRMQ

101 ILMIRYPSSQ LGEMSLEEHS QCECRPKKKD SAVKPDRAAT PHHRPQPRSV

151 PGWDSAPGAP SPADITHPTP APGPSAHAAP STTSALTPGP AAAAADAAAS

201 SVAKGGA

!!AA_SEQUENCE 1.0
ID R99253 standard; Protein; 374 AA.
AC R99253;
DT 05-DEC-1996 (first entry)
DE Cytoplasmic antiproteinase-2 protein.
KW Cytoplasmic antiproteinase-2 protein; CAP-2; serpin;
KW serine protease inhibitor; antiinflammatory; apoptosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT modified_site 8
FT /label= N-glycosylation_site
FT modified_site 78
FT /label= N-glycosylation_site
PN WO9624650-A2.
PD 15-AUG-1996.
PF 02-FEB-1996; U01288.
PR 08-FEB-1995; US-385500.
PA (ZYMO) ZYMOGENETICS INC.
PI Sprecher CA;
DR WPI; 96-393014/39.
DR N-PSDB; T35220.
PT Human cytoplasmic antiproteinase-2 (CAP-2) and CAP-3 - serine
PT protease inhibitors useful in the purification of proteins and in
PT the treatment of inflammatory diseases and apoptosis
PS Claim 9; Page 37-38; 50pp; English.
CC Human cytoplasmic antiproteinase-2 protein (CAP-2) (R99253) is an
CC inhibitor of specific trypsin-like serine proteases. Its amino
CC acid sequence was deduced from a cDNA clone (T35220) isolated from
CC a human placenta lambda gt11 cDNA library. Another serpin, CAP-3
CC (R99254), has also been identified. The CAPs are useful for
CC the purification of proteins and in the treatment of inflammatory
CC diseases and apoptosis. They can be expressed in transformed or
CC transfected host (pref. mammalian) cells.
SQ Sequence 374 AA;

R99253 Length: 374 February 11, 2000 15:48 Type: P Check: 4145 ..

1 MDDLCEANGT FAISLFKILG EEDNSRNVFF SPMSISSALA MVFMGAKGST

51 AAQMSQALCL YKDGDIHRGF QSLLSEVNRT GTQYLLRTAN RLFGEKTCDF

101 LPDFKEYCQK FYQAELEELS FAEDTEECRK HINDWVAEKT EGKISEVLDA

151 GTVDPLTKLV LVNAIYFKGK WNEQFDRKYT RGMLFKTNEE KKTVQMMFKE

201 AKFKMGYADE VHTQVLELPY VEEELSMVIL LPDDNTDLAV VEKALTYEKF

```
     251  KAWTNSEKLT KSKVQVFLPR LKLEESYDLE PFLRRLGMID AFDEAKADFS

     301  GMSTEKNVPL SKVAHKCFVE VNEEGTEAAA ATAVVRNSRC SRMEPRFCAD

     351  HPFLFFIRRH KTNCILFCGR FSSP

!!AA_SEQUENCE 1.0
ID   W01743 standard; Protein; 259 AA.
AC   W01743;
DT   22-APR-1997  (first entry)
DE   C. pneumoniae 53 kDa antigen.
KW   antigen; antibody; detection; determination; epitope.
OS   Chlamydia pneumoniae.
PN   J08304404-A.
PD   22-NOV-1996.
PF   28-APR-1995; 106014.
PR   28-APR-1995; JP-106014.
PA   (HITB ) HITACHI CHEM CO LTD.
DR   WPI; 97-056178/06.
DR   N-PSDB; T59311.
PT   Detection and determination of anti-Chlamydia pneumoniae antibody -
PT   using the polypeptide C as the antigen
PS   Example 3; Page 17-18; 18pp; Japanese.
CC   The protein is that of the Chlamydia pneumoniae 53 kDa antigen. A
CC   method for the detection and determination of anti-C. pneumoniae
CC   antibodies in a sample comprises using at least 5 consecutive amino
CC   acids of the polypeptide C 73 kDa antigen (W01742).
SQ   Sequence   259 AA;

 W01743  Length: 259  February 11, 2000 15:48  Type: P  Check: 9728  ..

       1  MSISSSSGPD NQKNIMSQVL TSTPQGVPQQ DKLSGNETKQ IQQTRQGKNT

      51  EMESDATIAG ASGKDKTSST TKTETAPQQG VAAGKESSES QKAGADTGVS

     101  GAAATTASNT ATKIAMQTSI EEASKSMEST LESLQSLSAA QMKEVEAVVV

     151  AALSGKSSGS AKLETPELPK PGVTPRSEVI EIGLALAKAI QTLGEATKSA

     201  LSNYASTQAQ ADQTNKLGLE KQAIKIDKER EEYQEMKAAE QKSKDLEGTM

     251  DTVNTVMIA

!!AA_SEQUENCE 1.0
ID   W04831 standard; Protein; 207 AA.
AC   W04831;
DT   28-APR-1997  (first entry)
DE   Vascular endothelial growth factor-B186.
KW   Endothelial cell; proliferation; vascular endothelial growth factor; VPF;
KW   VEGF; endothelium; mesodermal cell; cationic dimer; tissue regeneration;
KW   vascular permeability factor; cell mitogen; angiogenesis; cell growth;
KW   embryonic development; wound healing; tissue reorganisation; antibody;
KW   cancer; metastatic risk; tumour cell; mouse.
OS   Homo sapiens.
PN   WO9626736-A1.
PD   06-SEP-1996.
PF   01-MAR-1996; U02957.
PR   01-MAR-1995; US-397651.
PR   06-JUN-1995; US-469427.
PR   06-DEC-1995; US-569063.
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PI   Alitalo K,  Eriksson U,  Olofsson B,  Pajusola K;
DR   WPI; 96-412582/41.
DR   N-PSDB; T37915.
PT   Vascular endothelial growth factor VEGF-B proteins - useful to
PT   accelerate angiogenesis in wound healing, also related nucleic acid
PT   and antibodies for cancer diagnosis
PS   Claim 18; Page 62; 107pp; English.
CC   W04824-W04831 represent the vascular endothelial growth factor (VEGF)
CC   proteins of the invention, which promote endothelial or mesodermal cell
CC   proliferation.  VEGF is also a glycosylated cationic dimer, and is
CC   sometimes referred to as vascular permeability factor (VPF).  VEGF has
CC   diverse effects, depending on the specific biological context in which it
CC   is found.  VEGF is a potent endothelial cell mitogen, and directly
CC   contributes to induction of angiogenesis in vivo by promoting endothelial
CC   cell growth during normal embryonic development, wound healing, and
CC   tissue regeneration/reorganisation.  The VEGF proteins of the invention
CC   share the angiogenic and other properties of VEGF, but are distributed
CC   and expressed in tissues differently to VEGF.  The proteins can therefore
CC   be used to accelerate angiogenesis in wound healing.  Antibodies against
CC   the proteins can be used for inhibiting angiogenesis.  The antibodies can
CC   also be used diagnostically to quantitatively detect VEGF-B.  Primers
CC   complementary to the coding sequences for the proteins of the invention
CC   can also be used to detect VEGF-B coding sequences.  Quantification of
CC   VEGF-B in cancer biopsy specimens may be useful as an indicator of
CC   metastatic risk.  VEGF-B expression in a cell can be retarded using
CC   antisense sequences direct against the VEGF coding sequences, this is
CC   especially useful in retarding VEGF expression in tumour cells.
SQ   Sequence   207 AA;

 W04831  Length: 207  February 11, 2000 15:48  Type: P  Check: 1679  ..

       1  MSPLLRRLLL AALLQLAPAQ APVSQPDAPG HQRKVVSWID VYTRATCQPR

      51  EVVVPLTVEL MGTVAKQLVP SCVTVQRCGG CCPDDGLECV PTGQHQVRMQ

     101  ILMIRYPSSQ LGEMSLEEHS QCECRPKKKD SAVKPDRAAT PHHRPQPRSV

     151  PGWDSAPGAP SPADITHPTP APGPSAHAAP STTSALTPGP AAAAADAAAS

     201  SVAKGGA

!!AA_SEQUENCE 1.0
ID   W11481 standard; Protein; 205 AA.
AC   W11481;
DT   22-APR-1997  (first entry)
DE   D. immitis mature venom allergen antigen 5-like protein PDiVA205.
KW   Venom allergen antigen 5-like gene; VA5; helminth; parasite;
KW   nDiVA615; PDiVA205; vaccine; heartworm.
OS   Dirofilaria immitis.
PN   WO9637218-A1.
PD   28-NOV-1996.
PF   23-MAY-1996; U07709.
PR   23-MAY-1995; US-450944.
PA   (HESK-) HESKA CORP.
PI   Tripp CA,  Wisnewski N;
DR   WPI; 97-020935/02.
DR   N-PSDB; T51379.
PT   Nucleic acid encoding helminth venom allergen antigen 5-like protein
PT   - pref. from Dirofilaria immitis or Onchocerca volvulus, useful esp.
PT   in vaccines to prevent helminth infection
PS   Claim 19; Page 91-92; 124pp; English.
CC   Dirofilaria immitis mature venom allergen antigen 5-like protein
CC   VA5 (W11481), or PDiVA205, is capable of eliciting an immune
CC   response (cellular and/or humoral) in animals.  It is encoded by a
CC   cDNA molecule nDiVA615 (T51379).  VA5-like proteins (see also
CC   W11480-86) can be used in vaccines to prevent helminth infection.
SQ   Sequence   205 AA;

 W11481  Length: 205  February 11, 2000 15:48  Type: P  Check: 5195  ..

       1  YECEGGKLTP TERKNIVTQI NKYRSRLIRG KLKNKDGYLM PKGKNMLRMR

      51  WDCKLEKSAQ NWANMCVFGH SPSSERRGIG ENVYAYWSSG SVRDLKKTAG

     101  TDAGRLWWSE LEKYYSDNPS NNLTSEVAME NILHFTQMAW GETYKLGSGV

     151  DHNIVMVART LVFICHYFPG GNMVKDLIYE LGNPCKHNKD CRTKRCSAKS

     201  GLCKK

!!AA_SEQUENCE 1.0
ID   W11479 standard; Protein; 221 AA.
AC   W11479;
DT   22-APR-1997  (first entry)
```

```
DE   D. immitis venom allergen antigen 5-like protein PDiVA221.
KW   Venom allergen antigen 5-like gene; VA5; helminth; parasite;
KW   nDiVA833; PDiVA221; vaccine; heartworm.
OS   Dirofilaria immitis.
FH   Key             Location/Qualifiers
FT   peptide         1. .16
FT                   /label= Sig_peptide
FT   protein         17. .221
FT                   /label= Mat_protein
FT                   /note= "mature protein is represented by PDiVA205"
PN   WO9637218-A1.
PD   28-NOV-1996.
PF   23-MAY-1996; U07709.
PR   23-MAY-1995; US-450944.
PA   (HESK-) HESKA CORP.
PI   Tripp CA,  Wisnewski N;
DR   WPI; 97-020935/02.
DR   N-PSDB; T51376;
DR   N-PSDB; T51378.
PT   Nucleic acid encoding helminth venom allergen antigen 5-like protein
PT   - pref. from Dirofilaria immitis or Onchocerca volvulus, useful esp.
PT   in vaccines to prevent helminth infection
PS   Claim 13; Page 89-90; 124pp; English.
CC   Dirofilaria immitis venom allergen antigen 5-like protein VA5
CC   (W11479), or PDiVA221, is capable of eliciting an immune response
CC   (cellular and/or humoral) in animals.  It is encoded by a cDNA
CC   clone (T51376) isolated from a D. immitis cDNA library by screening
CC   with immune dog serum.  VA5-like proteins (see also W11480-86) can
CC   be used in vaccines to prevent helminth infection.
SQ   Sequence   221 AA;

 W11479  Length: 221  February 11, 2000 15:48  Type: P  Check: 9389  ..

       1  MILLVIFPAI IVVVASYECE GGKLTPTERK NIVTQINKYR SRLIRGKLKN

      51  KDGYLMPKGK NMLRMRWDCK LEKSAQNWAN MCVFGHSPSS ERRGIGENVY

     101  AYWSSGSVRD LKKTAGTDAG RLWWSELEKY YSDNPSNNLT SEVAMENILH

     151  FTQMAWGETY KLGSGVDHNI VMVARTLVFI CHYFPGGNMV KDLIYELGNP

     201  CKHNKDCRTK RCSAKSGLCK K

!!AA_SEQUENCE 1.0
ID   W13272 standard; Protein; 244 AA.
AC   W13272;
DT   02-JUN-1997  (first entry)
DE   Rhodococcus erythropolis SK92-B1 regulatory factor component.
KW   Regulatory factor; activation; nitrilase; gene promoter;
KW   recombinant plasmid; transformation; microorganism; catalyst;
KW   production; organic acid.
OS   Rhodococcus erythropolis.
PN   J09028382-A.
PD   04-FEB-1997.
PF   19-JUL-1995; 204061.
PR   19-JUL-1995; JP-204061.
PA   (NITT ) NITTO CHEM IND CO LTD.
DR   WPI; 97-159092/15.
DR   N-PSDB; T46931.
PT   Microorganism producing nitrilase in absence of inducer - useful as
PT   catalyst in large scale organic acid production
PS   Claim 7; Pages 7-8; 13pp; Japanese.
CC   The present sequence is one of the two components of the
CC   Rhodococcus erythropolis SK92-B1 regulatory factor (RF), which
CC   activates a nitrilase gene promoter. A recombinant plasmid
CC   containing the RF gene can be used to transform a microorganism,
CC   enabling it to produce nitrilase in high yield without adding any
CC   inducer to the medium. Hence, the preparation of the medium is easy
CC   for industrial large scale culture of the microorganism, and the
CC   disposal cost of the waste culture is reduced. Nitrilase is
CC   industrially useful as a catalyst for the production of organic
CC   acids, as it hydrolyses nitriles into their corresponding acids.
SQ   Sequence   244 AA;

 W13272  Length: 244  February 11, 2000 15:48  Type: P  Check: 2018  ..

       1  MAGADVHAQG GTNRRARILV VDDEKHVRTM VTWQLESENF DVVAAADGDA

      51  ALRQVTESAP DLMVLDLSLP GKGGLEVLAT VRRTDALPIV VLTARRDETE

     101  RIVALDLGAD DYVIKPFSPR ELAARIRAVL RRTTAEPPHE AAVQRFGDLE

     151  IDTAAREVRL HGIPLEFTTK EFDLLAYMAA SPMQVFSRRR LLLEVWRSSP

     201  DWQQDATVTE HVHRIRRKIE EDPTKPTILQ TVRGAGYRFD GERA

!!AA_SEQUENCE 1.0
ID   W06943 standard; Protein; 140 AA.
AC   W06943;
DT   30-JUN-1997  (first entry)
DE   CagI locus product 14.
KW   CagI; CagA; virulence factor; exporter molecule; homology; ptl gene;
KW   Bordetella pertussis; VIR B4; Agrobacterium tumefaciens; invasion factor;
KW   Salmonella; type I strain; virulence; diagnosis; H. pylori; infection;
KW   vaccine; treatment; duodenal; gastric ulcer; active gastritis;
KW   adenocarcinoma.
OS   Helicobacter pylori.
PN   WO9633274-A1.
PD   24-OCT-1996.
PF   18-APR-1996; IB0343.
PR   20-APR-1995; US-425194.
PR   07-JUN-1995; US-477451.
PA   (BIOC-) BIOCINE SPA.
PI   Covacci A;
DR   WPI; 96-485780/48.
DR   N-PSDB; T46159.
PT   Helicobacter pylori CagI polynucleotide and related proteins - used
PT   in diagnosis and in vaccines for the treatment of H. pylori
PT   infection associated disease
PS   Disclosure; Fig 4I; 303pp; English.
CC   The present sequence is encoded by a putative open reading frame (ORF) of
CC   the CagI locus (see T46159). The CagI region contains clusters of
CC   putative ORFs with different polarities. It is hypothesised that some of
CC   the ORFs may encode exporter molecules with homology to the ptl genes of
CC   Bordetella pertussis and VIR B4 genes of Agrobacterium tumefaciens and
CC   for proteins with motifs shared by the purported invasion factors of
CC   Salmonella genus. The absence of the CagA gene in the type I strains is
CC   associated with the absence of CagI sequences (which may encode virulence
CC   factors restricted to type I strains). The CagI nucleotide sequence, its
CC   fragments and encoded proteins are used in the diagnosis of H. pylori
CC   (esp. H. pylori type I strain) infection in an individual and in vaccines
CC   (claimed) for the treatment of H. pylori infection associated with e.g.
CC   duodenal and gastric ulcers, severe forms of active gastritis (esp. type
CC   gastritis) and gastric adenocarcinoma.
SQ   Sequence   140 AA;

 W06943  Length: 140  February 11, 2000 15:48  Type: P  Check: 9316  ..

       1  KRIFIKLNYY LLGVLSLACL QSVNADQNTD IKDISPEDMA LNSVGLVSRD

      51  QLKIEIPKET LEQKVAILND YNDKNVNIKF DDISLGSFQP NDNLGINAMW

     101  GIQNLLMSQM MGDYGPNNPF MYGYAPTYSD SSFLPPILGY

!!AA_SEQUENCE 1.0
ID   W06935 standard; Protein; 136 AA.
AC   W06935;
DT   30-JUN-1997  (first entry)
DE   CagI locus product 6.
KW   CagI; CagA; virulence factor; exporter molecule; homology; ptl gene;
KW   Bordetella pertussis; VIR B4; Agrobacterium tumefaciens; invasion factor;
KW   Salmonella; type I strain; virulence; diagnosis; H. pylori; infection;
KW   vaccine; treatment; duodenal; gastric ulcer; active gastritis;
KW   adenocarcinoma.
OS   Helicobacter pylori.
PN   WO9633274-A1.
```

PD 24-OCT-1996.
PF 18-APR-1996; IB0343.
PR 20-APR-1995; US-425194.
PR 07-JUN-1995; US-477451.
PA (BIOC-) BIOCINE SPA.
PI Covacci A;
DR WPI; 96-485780/48.
DR N-PSDB; T46159.
PT Helicobacter pylori CagI polynucleotide and related proteins - used
PT in diagnosis and in vaccines for the treatment of H. pylori
PT infection associated disease
PS Disclosure; Fig 4D; 303pp; English.
CC The present sequence is encoded by a putative open reading frame (ORF) of
CC the CagI locus (see T46159). The CagI region contains clusters of
CC putative ORFs with different polarities. It is hypothesised that some of
CC the ORFs may encode exporter molecules with homology to the ptl genes of
CC Bordetella pertussis and VIR B4 genes of Agrobacterium tumefaciens and
CC for proteins with motifs shared by the purported invasion factors of
CC Salmonella genus. The absence of the CagA gene in the type I strains is
CC associated with the absence of CagI sequences (which may encode virulence
CC factors restricted to type I strains). The CagI nucleotide sequence, its
CC fragments and encoded proteins are used in the diagnosis of H. pylori
CC (esp. H. pylori type I strain) infection in an individual and in vaccines
CC (claimed) for the treatment of H. pylori infection associated with e.g.
CC duodenal and gastric ulcers, severe forms of active gastritis (esp. type
CC gastritis) and gastric adenocarcinoma.
SQ Sequence 136 AA;

W06935 Length: 136 February 11, 2000 15:48 Type: P Check: 8017 ..

1 NIYYFMLLYK FTALNYFFKN QNGTNQISKL KQNFLQFKYS FNKHLDKYSL

51 YYRLFNISSI VIGFLIGLFS YGAGVILVYP ILFLFALIIK PSFFYYTTYL

101 LLLVSLSIIS KYYLLSHAKF TMKLIILMTQ WQNWFL

!!AA_SEQUENCE 1.0
ID W20571 standard; Protein; 114 AA.
AC W20571;
DT 17-JUL-1997 (first entry)
DE H. pylori secreted or periplasmic protein 80257.aa.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL;
DR WPI; 97-052306/05.
DR N-PSDB; T67714.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 72; Page 725-726; 1481pp; English.
CC This sequence is a H. pylori secreted or periplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 114 AA;

W20571 Length: 114 February 11, 2000 15:48 Type: P Check: 6963 ..

1 MKRPISKLKQ NFLQFKHSFN KHLDKYSLYY RLFNISSIVI GFLIALFSYG

51 AGVILVYPIL FLFALIIKPS FFYYTTYLLL LVSLSIISKY YLLSHANFTM

101 KLIMLMTQWQ NWFL

!!AA_SEQUENCE 1.0
ID W20886 standard; Protein; 154 AA.
AC W20886;
DT 18-JUL-1997 (first entry)
DE H. pylori secreted or periplasmic protein, 13eel2016orf8.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacterium; life cycle; activator;
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW diagnosis.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL;
DR WPI; 97-052306/05.
DR N-PSDB; T68139.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 72; Page 1289; 1481pp; English.
CC This sequence represents a H. pylori secreted or periplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 154 AA;

W20886 Length: 154 February 11, 2000 15:48 Type: P Check: 2872 ..

1 NNAKNKNKGV KSMKTNFYKI KLLFAWCLII GMFNAPLNAD QNTDIKDISP

51 EDMALNSVGL VSRDQLKIEI PKETLEQKVA ILNDYNDKNV NIKFDDISLG

101 SFQPNDNLGI NAMWGIQNLL MSQMMSNYGP NNSFMYGYAP TYSDSSFLPP

151 ILGY

!!AA_SEQUENCE 1.0
ID W20776 standard; Protein; 276 AA.
AC W20776;
DT 15-JUL-1997 (first entry)
DE H. pylori flagella-associated protein, 07ge20415orf34.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacterium; life cycle; activator;
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW diagnosis.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.

PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL;
DR WPI; 97-052306/05.
DR N-PSDB; T68029.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 56; Page 1187; 1481pp; English.
CC This sequence represents a H. pylori flagella-associated protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 276 AA;

W20776 Length: 276 February 11, 2000 15:48 Type: P Check: 6185 ..

1 RLKMATKLTP KQKAQLDELS MSEKIAILLI QVGEDTTGEI LRHLDIDSIT

51 EISKQIVQLN GTDKQIGAAV LEEFFAIFQS NQYINTGGLE YARELLTRTL

101 GSEEARKVMD KLTKSLQTQK NFAYLGKIKP QQLADFIINE HPQTIALILA

151 HMEAPNAAET LSYFPDEMKA EISIRMANLG EISPQVVKRV STVLENKLES

201 LTSYKIEVGG LRAVAEIFNR LGQKSAKTTL ARIESVDNKL AGAIKEMMFT

251 FEDIAKLDNF AIMRDFKSGG LKKTGL

!!AA_SEQUENCE 1.0
ID W20259 standard; Protein; 108 AA.
AC W20259;
DT 09-JUL-1997 (first entry)
DE H. pylori secreted or periplasmic protein 23564012.aa.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL;
DR WPI; 97-052306/05.
DR N-PSDB; T67473.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 72; Page 463; 1481pp; English.
CC This sequence is a H. pylori secreted or periplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 108 AA;

W20259 Length: 108 February 11, 2000 15:48 Type: P Check: 5141 ..

1 MKTNFYKIKL LFAWCLIIGM FNAPLNADQN TDIKDISPED MALNSVGLVS

51 RDQLKIEIPK ETLEQKVTIL NDYNDKNVNI KFDDISLGSF QPNDNLGINA

101 MWGIQNLL

!!AA_SEQUENCE 1.0
ID W20324 standard; protein; 102 AA.
AC W20324;
DT 09-JUL-1997 (first entry)
DE H. pylori cell envelope transporter protein, 2461062.aa.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaerd BL;
DR WPI; 97-052306/05.
DR N-PSDB; T67528.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 56; Pages 516; 1481pp; English.
CC The present sequence is a Helicobacter pylori cell envelope protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds, useful
CC as potential H. pylori life cycle activators or inhibitors. The genomic
CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF of at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 102 AA;

W20324 Length: 102 February 11, 2000 15:48 Type: P Check: 6372 ..

1 MMAHSLILVS KTSLSNLLIF VVQPDGKLSM TDAAIDPNMT NSGLRWYRVN

51 EIAEKFKLIK DKALVTVINK GYGKNPLTKN YNIKNYGELE RVIKKLPLVR

101 DK

!!AA_SEQUENCE 1.0
ID W20696 standard; Protein; 121 AA.
AC W20696;
DT 15-JUL-1997 (first entry)
DE H. pylori secreted or periplasmic protein 05ae20220orf50.
KW Secretion; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;
KW periplasm.
OS Helicobacter pylori.
PN WO9640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR) ASTRA AB.

```
PI   Berglindh OT,  Smith D,  Mellgaerd BL;
DR   WPI; 97-052306/05.
DR   N-PSDB; T67949.
PT   Helicobacter pylori nucleic acid sequences and related
PT   polypeptide(s)  - useful for vaccines to treat or prevent H. pylori
PT   infection, and to detect Helicobacter
PS   Claim 72; Page 1115; 1481pp; English.
CC   The present sequence represents a Helicobacter pylori protein which
CC   is likely to be secreted or periplasmic.
CC   The protein may be used in a vaccine to prevent or treat H. pylori
CC   infection or to identify H. pylori polypeptide binding compounds,
CC   useful as potential H. pylori life cycle activators or inhibitors.
CC   The genomic sequence of H. pylori (ATCC 55679) was determined from
CC   overlapping contigs generated by mechanically shearing the bacterial
CC   DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC   and the predicted coding regions defined by computer evaluation. To
CC   identify likely H. pylori antigens for vaccine development, the amino
CC   acid sequences predicted from various ORF were analysed for significant
CC   homology to other known or exported membrane proteins. Having identified
CC   and determined the sequences of interest, particular regions can be
CC   isolated from H. pylori by PCR amplification for recombinant polypeptide
CC   production, e.g. in E. coli hosts.
SQ   Sequence   121 AA;

 W20696  Length: 121  February 11, 2000 15:48  Type: P  Check: 7893  ..

       1  TSSFKTKMKR PISKLKQNFL QFKHSFNKHL DKYSLYYRLF NISSIVIGFL

      51  IALFSYGAGV ILVYPILFLF ALIIKPSFFY YTTYLLLLVS LSIISKYYLL

     101  SHANFTMKLI MLMTQWQNWF L

!!AA_SEQUENCE 1.0
ID   W20139 standard; protein; 53 AA.
AC   W20139;
DT   22-JUL-1997  (first entry)
DE   H. pylori cytoplasmic protein, 14257751.aa.
KW   Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW   binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW   duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS   Helicobacter pylori.
PN   WO9640893-A1.
PD   19-DEC-1996.
PF   06-JUN-1996; U09122.
PR   07-JUN-1995; US-487032.
PR   01-APR-1996; US-630405.
PA   (ASTR ) ASTRA AB.
PI   Berglindh OT,  Smith D,  Mellgaerd BL;
DR   WPI; 97-052306/05.
DR   N-PSDB; T67382.
PT   Helicobacter pylori nucleic acid sequences and related
PT   polypeptide(s)  - useful for vaccines to treat or prevent H. pylori
PT   infection, and to detect Helicobacter
PS   Claim 61; Page 363; 1481pp; English.
CC   The present sequence is a H. pylori cytoplasmic protein involved in
CC   carbohydrate metabolism.
CC   The protein may be used in a vaccine to prevent or treat H. pylori
CC   infection or to identify H. pylori polypeptide binding compounds,
CC   useful as potential H. pylori life cycle activators or inhibitors.
CC   The genomic sequence of H. pylori (ATCC 55679) was determined from
CC   overlapping contigs generated by mechanically shearing the bacterial
CC   DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC   and the predicted coding regions defined by computer evaluation. To
CC   identify likely H. pylori antigens for vaccine development, the amino
CC   acid sequences predicted from various ORF were analysed for significant
CC   homology to other known or exported membrane proteins. Having identified
CC   and determined the sequences of interest, particular regions can be
CC   isolated from H. pylori by PCR amplification for recombinant polypeptide
CC   production, e.g. in E. coli hosts.
SQ   Sequence   53 AA;

 W20139  Length: 53  February 11, 2000 15:48  Type: P  Check: 1367  ..

       1  MLDFDLVLFG ATGDLAMRKL FVSLYEIYIS FMVLKTILGL SHRGVRSYPM

      51  KSF

!!AA_SEQUENCE 1.0
ID   W09433 standard; Protein; 328 AA.
AC   W09433;
DT   01-SEP-1997  (first entry)
DE   Human placenta purinergic P-2u receptor, PNR.
KW   PNR; purinergic P-2u receptor; placental; inflammation; diagnosis;
KW   carcinoma; neoplasia; cancer; cystic fibrosis; hypertension;
KW   high blood pressure; infection.
OS   Homo sapiens.
PN   WO9638558-A2.
PD   05-DEC-1996.
PF   03-JUN-1996; U08481.
PR   02-JUN-1995; US-459046.
PA   (INCY-) INCYTE PHARM INC.
PI   Au-Young J,  Coleman R,  Stuart SG;
DR   WPI; 97-042714/04.
DR   N-PSDB; T47730.
PT   New isolated human placenta purinergic P-2u receptor (PNR) gene -
PT   used to develop prods. for the diagnosis and treatment of conditions
PT   associated with altered expression of the receptor e.g. inflammation
PS   Claim 1; Page 27-29; 36pp; English.
CC   W09433 shows a human placenta-derived purinergic P-2u receptor
CC   designated PNR. P-2u is specifically expressed in cells active in
CC   immunity. An assay for up-regulated expression of PNR can accelerate
CC   diagnosis and proper treatment of conditions caused by abnormal signal
CC   transduction due to systemic and local infections, traumatic and
CC   other tissue damage, hereditary or environmental diseases associated
CC   with hypertension, carcinomas, cystic fibrosis and other physiological
CC   or pathological problems.
SQ   Sequence   328 AA;

 W09433  Length: 328  February 11, 2000 15:48  Type: P  Check: 8653  ..

       1  MEWDNGTDQA LGLPPTTCVY RENFKQLLLP PVYSAVLAPA LPLNICVITQ

      51  ICTSRRALTR TAVYTLNLAL PDLLYACSLP LLIYNYAQGD HWPFGDFACR

     101  LVRFLFYANL HGRILFLTCI SFQRYLGICH PLAPWHKRGG RRAAWLVCVA

     151  VWLAVTTQCL PTAIFAATGI QRNRTVCYDL SPPALATHYM PYGMALTVIG

     201  FLLPFAALLA CYCLLACRLC RQDGPAEPVA QERRGKAARM AVVVAAVFGI

     251  SFLPFHITKT AYLAVRSTPG VPCTVLEAFA AAYKGTRPFA SANSVLDPIL

     301  FYFTQKKFRR RPHELLQKLT DKWQRQGR

!!AA_SEQUENCE 1.0
ID   W21704 standard; Protein; 248 AA.
AC   W21704;
DT   26-SEP-1997  (first entry)
DE   Luffin-A.
KW   pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;
KW   inactivation; eukaryotic ribosome; alpha fragment; beta fragment;
KW   inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;
KW   rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.
OS   Luffa cylindrica.
FH   Key             Location/Qualifiers
FT   region          138. .148
FT                   /note= "Position of possible insertion of internal
FT                   peptide linker sequence"
PN   US5635384-A.
PD   03-JUN-1997.
PF   11-JUN-1990; 535636.
PR   11-JUN-1990; US-535636.
PR   09-DEC-1992; US-987927.
PR   26-JAN-1995; US-378761.
PA   (DOWC ) DOWELANCO.
PI   Hey TD,  Morgan AER,  Walsh TA;
```

DR WPI; 97-309831/28.
PT Inactive precursor of maize ribosome-inactivating protein - also
PT chimeric ribosome-inactivating protein precursors containing
PT internal linker sequences
PS Claim 2; Column 117-118; 121pp; English.
CC The sequences given in W21698-710 represent Ribosome Inactivating
CC Proteins (RIP's), which may be used in the construction of the
CC proRIP of the invention. The proRIP has a selectively removable,
CC internal peptide linker. The precursor sequence is incapable of
CC inactivating eukaryotic ribosomes, but can be converted by removal
CC of the linker into a protein having alpha and beta fragments and being
CC capable of inactivating eukaryotic ribosomes. RIPs are potent
CC inhibitors of eukaryotic protein synthesis. They possess a highly
CC specific N-glycosidase activity which cleaves the glycosidic bond of
CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit
CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T
CC cells. The inactive proRIP proteins make it possible to provide protein
CC synthesis inhibitors with uses in practical and improved ways not before
CC possible. The RIP can be used to make cytotoxic conjugates.
SQ Sequence 248 AA;

W21704 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..

1 DVRFSLSGSS STSYSKFIGD LRKALPSNGT VYNLTILLSS ASGASRYTLM

51 TLSNYDGKAI TVAVDVSQLY IMGYLVNSTS YFFNESDAKL ASQYVFKGST

101 IVTLPYSGNY EKLQTAAGKI REKIPLGFPA LDSALTTIFH YDSTAAAAAF

151 LVILQTTAEA SRFKYIEGQI IERISKNQVP SLATISLENS LWSALSKQIQ

201 LAQTNNGTFK TPVVITDDKG QRVEITNVTS KVVTKNIQLL LNYKQNVA

!!AA_SEQUENCE 1.0
ID W21709 standard; Protein; 250 AA.
AC W21709;
DT 26-SEP-1997 (first entry)
DE Luffin-B.
KW pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;
KW inactivation; eukaryotic ribosome; alpha fragment; beta fragment;
KW inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;
KW rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.
OS Luffa cylindrica.
FH Key Location/Qualifiers
FT region 139. .149
FT /note= "Position of possible insertion of internal
FT peptide linker sequence"
PN US5635384-A.
PD 03-JUN-1997.
PF 11-JUN-1990; 535636.
PR 11-JUN-1990; US-535636.
PR 09-DEC-1992; US-987927.
PR 26-JAN-1995; US-378761.
PA (DOWC) DOWELANCO.
PI Hey TD, Morgan AER, Walsh TA;
DR WPI; 97-309831/28.
PT Inactive precursor of maize ribosome-inactivating protein - also
PT chimeric ribosome-inactivating protein precursors containing
PT internal linker sequences
PS Claim 2; Column 125-128; 121pp; English.
CC The sequences given in W21698-710 represent Ribosome Inactivating
CC Proteins (RIP's), which may be used in the construction of the
CC proRIP of the invention. The proRIP has a selectively removable,
CC internal peptide linker. The precursor sequence is incapable of
CC inactivating eukaryotic ribosomes, but can be converted by removal
CC of the linker into a protein having alpha and beta fragments and being
CC capable of inactivating eukaryotic ribosomes. RIPs are potent
CC inhibitors of eukaryotic protein synthesis. They possess a highly
CC specific N-glycosidase activity which cleaves the glycosidic bond of
CC adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit
CC cellular proliferation of cells, e.g. cancer cells and HIV-infected T
CC cells. The inactive proRIP proteins make it possible to provide protein
CC synthesis inhibitors with uses in practical and improved ways not before
CC possible. The RIP can be used to make cytotoxic conjugates.
SQ Sequence 250 AA;

W21709 Length: 250 February 11, 2000 15:48 Type: P Check: 7681 ..

1 ANVSFSLSGA DSKSYSKFIT ALRKALPSKE KVSNIPLLLP SASGASRYIL

51 MQLSNYDAKA ITMAIDVTNV YIMGYLVNST SYFANESDAK LASQYVFKGS

101 TLVTIPYSGN YERLQNAAGK IREKIPLGFR ALDSALTSIF HYDSTAAAAA

151 FLVILQTTAE ASRFKYIEGQ IIERIPKNEV PSPAALSLEN EAWSLLSKQI

201 QLAQTNNGAF RTPVVIIDNK GQRVEITNLA SKVQIKDVNS KLLLNKQNIA

!!AA_SEQUENCE 1.0
ID W10031 standard; Protein; 74 AA.
AC W10031;
DT 24-SEP-1997 (first entry)
DE Protein encoded by clone G-4-5 from activated maize embryo.
KW primer; polymerase chain reaction; cDNA library; anchored end; PCSUB;
KW lock-docking; screening; PCR-based cDNA subtractive cloning; Zea mays;
KW maize; Fusarium moniliforme.
OS Zea mays.
PN WO9640998-A1.
PD 19-DEC-1996.
PF 05-JUN-1996; U08582.
PR 07-JUN-1995; US-481687.
PA (PION-) PIONEER HI-BRED INT INC.
PI Briggs SP, Duvick JP, Wang X;
DR WPI; 97-087067/08.
DR N-PSDB; T70121.
PT Method for prodn. of cDNA libraries with anchored ends - useful for
PT subtractive cloning of sequences of interest
PS Disclosure; Fig 9; 56pp; English.
CC This protein is the longest putative open reading frame encoded by cDNA
CC clone G-4-5 which was activated upon Fusarium moniliforme infection of
CC germinating maize embryos. This cDNA clone was identified from a full-
CC length cDNA library with anchored ends using a method of isolation
CC designated ICEFAL (isolation of cDNA ends from an anchored library). The
CC method of generating a cDNA library with anchored ends produces discreet
CC sized PCR products which do not necessarily require further subcloning/
CC screening. The method also produces full-length cDNA's obtainable from
CC the libraries as opposed to specific cDNA clones, as produced by
CC previously known methods. Other methods such as PCR and RACE require a
CC knowledge of the target sequence to be amplified, by using the PCSUB
CC method no previous knowledge is necessary.
SQ Sequence 74 AA;

W10031 Length: 74 February 11, 2000 15:48 Type: P Check: 8373 ..

1 WILRRLPHVH PPQGASREEP GRLPQGRRRR RQTVLLASIY LASLLPIEIV

51 CMCGVCMHLP TLTTVHMYAA SCRP

!!AA_SEQUENCE 1.0
ID W13647 standard; peptide; 30 AA.
AC W13647;
DT 25-NOV-1997 (first entry)
DE Invariant chain region which contains the CLIP sequence.
KW Class II associated invariant chain peptide; CLIP region; MHC;
KW major histocompatibility complex; T cell epitope; antigen processing;
KW p35 isoform; vaccine; autoimmune disorder; rheumatoid arthritis;
KW multiple sclerosis; insulin-dependent diabetes mellitus;
KW myasthenia gravis; allergy; tumour; virus; parasite; mycobacterium;
KW infection; transplant rejection; immunostimulant; diagnosis.
OS Homo sapiens.
PN WO9710335-A1.
PD 20-MAR-1997.
PF 16-SEP-1996; D01763.
PR 14-MAY-1996; US-017663.
PR 14-SEP-1995; DE-034170.
PA (MELM/) MELMS A.

PA (UYTU-) UNIV TUEBINGEN.
PI Malcherek G, Melms A;
DR WPI; 97-202231/18.
PT Recombinant polypeptide useful as immunostimulant and diagnostic
PT tool - based on primary sequence of invariant chain contg. primary
PT sequence of specific T-cell epitope
PS Disclosure; Page 8; 29pp; German.
CC This is the sequence of amino acids 97-126 from the p35 isoform of
CC the invariant chain (i.e. of a polypeptide which associates with
CC major histocompatibility complex class II molecules in the
CC endoplasmic reticulum to form a class II loading compartment).
CC Claimed polypeptides are derived from the invariant chain by
CC replacing at least part of the CLIP (CLass II associated Invariant
CC chain Peptide) sequence with a T cell epitope. The recombinant
CC polypeptide is processed in antigen presenting cells to produce the
CC specific T cell epitope. Such polypeptides are useful in vaccines
CC for the treatment of autoimmune disorders, e.g. rheumatoid arthritis,
CC multiple sclerosis, insulin-dependent diabetes mellitus and myasthenia
CC gravis, allergies and tumours, for the control of viral, parasitic and
CC mycobacterial infections, and for use against transplant rejections.
CC They can also be used in the diagnosis of autoimmune disorders.
CC Vaccines produced using the recombinant polypeptides have a higher
CC immunostimulant specificity and selectivity over prior art vaccines.
SQ Sequence 30 AA;

W13647 Length: 30 February 11, 2000 15:48 Type: P Check: 5362 ..

1 LPKPPKPVSK MRMATPLLMQ ALPMGALPQG

!!AA_SEQUENCE 1.0
ID W25141 standard; Protein; 248 AA.
AC W25141;
DT 02-DEC-1997 (first entry)
DE Luffin-A (a ribosome inhibitory protein) inactive precursor.
KW Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;
KW internal linker; Barley Translation Inhibitor; Trichosanthin;
KW Ricin A-chain; Abrin-A A-chain; Saporin; SLT-1; Luffin A; MAP;
KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
KW post-translational modification; cancer; neoplasia; HIV: AIDS;
KW human immunodeficiency virus; acquired immune deficiency syndrome.
OS Synthetic.
PN US5646026-A.
PD 08-JUL-1997.
PF 11-JUN-1990; 535636.
PR 09-DEC-1992; US-987927.
PR 11-JUN-1990; US-535636.
PR 26-JAN-1995; US-378761.
PR 07-JUN-1995; US-485286.
PA (DOWC) DOWELANCO.
PI Hey TD, Morgan AER, Walsh TA;
DR WPI; 97-362934/33.
PT DNA encoding pro-ribosome inactivating proteins - inactive
PT precursors of ribosome inactivating proteins; can be expressed in
PT eukaryotic cells without causing cell death
PS Claim 4; Column 117-120; 186pp; English.
CC W25141 shows a Luffin-A (ribosome inhibitory protein, RIP) protein
CC which was engineered to contain a selectively removable internal peptide
CC linker sequence separating the alpha and beta units of the RIP. When
CC separated the two units regain activity and are capable of inactivating
CC eukaryotic ribosomes and hence preventing protein production. Many
CC different RIPs may be produced with an internal linker including
CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and
CC Saporin. The RIPs can be used in the construction of therapeutic
CC toxins targeted to specific cells such as tumour cells via the
CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.
CC A further use is in HIV therapy (see US4869903). There is interest
CC in expressing RIP recombinantly in host eukaryotic cells, because of
CC the capacity to provide correct post-translational processing. However,
CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting
CC in cell death. Since the inactive RIP proteins are not cytotoxic to
CC eukaryotic cells, they can be recombinantly expressed in such cells and
CC then converted to active RIP proteins.
SQ Sequence 248 AA;

W25141 Length: 248 February 11, 2000 15:48 Type: P Check: 7673 ..

1 DVRFSLSGSS STSYSKFIGD LRKALPSNGT VYNLTILLSS ASGASRYTLM

51 TLSNYDGKAI TVAVDVSQLY IMGYLVNSTS YFFNESDAKL ASQYVFKGST

101 IVTLPYSGNY EKLQTAAGKI REKIPLGFPA LDSALTTIFH YDSTAAAAAF

151 LVILQTTAEA SRFKYIEGQI IERISKNQVP SLATISLENS LWSALSKQIQ

201 LAQTNNGTFK TPVVITDDKG QRVEITNVTS KVVTKNIQLL LNYKQNVA

!!AA_SEQUENCE 1.0
ID W25146 standard; Protein; 250 AA.
AC W25146;
DT 02-DEC-1997 (first entry)
DE Luffin-B (a ribosome inhibitory protein) inactive precursor.
KW Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;
KW internal linker; Barley Translation Inhibitor; Trichosanthin;
KW Ricin A-chain; Abrin-A A-chain; Saporin; SLT-1; Luffin A; MAP;
KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
KW post-translational modification; cancer; neoplasia; HIV: AIDS;
KW human immunodeficiency virus; acquired immune deficiency syndrome.
OS Synthetic.
PN US5646026-A.
PD 08-JUL-1997.
PF 11-JUN-1990; 535636.
PR 09-DEC-1992; US-987927.
PR 11-JUN-1990; US-535636.
PR 26-JAN-1995; US-378761.
PR 07-JUN-1995; US-485286.
PA (DOWC) DOWELANCO.
PI Hey TD, Morgan AER, Walsh TA;
DR WPI; 97-362934/33.
PT DNA encoding pro-ribosome inactivating proteins - inactive
PT precursors of ribosome inactivating proteins; can be expressed in
PT eukaryotic cells without causing cell death
PS Claim 4; Column 127-128; 186pp; English.
CC W25146 shows a Luffin-B (ribosome inhibitory protein, RIP) protein
CC which was engineered to contain a selectively removable internal peptide
CC linker sequence separating the alpha and beta units of the RIP. When
CC separated the two units regain activity and are capable of inactivating
CC eukaryotic ribosomes and hence preventing protein production. Many
CC different RIPs may be produced with an internal linker including
CC maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and
CC Saporin. The RIPs can be used in the construction of therapeutic
CC toxins targeted to specific cells such as tumour cells via the
CC attachment of a targeting polypeptide, e.g. a monoclonal antibody.
CC A further use is in HIV therapy (see US4869903). There is interest
CC in expressing RIP recombinantly in host eukaryotic cells, because of
CC the capacity to provide correct post-translational processing. However,
CC RIPs effectively inhibit protein synthesis in eukaryotic cells resulting
CC in cell death. Since the inactive RIP proteins are not cytotoxic to
CC eukaryotic cells, they can be recombinantly expressed in such cells and
CC then converted to active RIP proteins.
SQ Sequence 250 AA;

W25146 Length: 250 February 11, 2000 15:48 Type: P Check: 7681 ..

1 ANVSFSLSGA DSKSYSKFIT ALRKALPSKE KVSNIPLLLP SASGASRYIL

51 MQLSNYDAKA ITMAIDVTNV YIMGYLVNST SYFANESDAK LASQYVFKGS

101 TLVTIPYSGN YERLQNAAGK IREKIPLGFR ALDSALTSIF HYDSTAAAAA

151 FLVILQTTAE ASRFKYIEGQ IIERIPKNEV PSPAALSLEN EAWSLLSKQI

201 QLAQTNNGAF RTPVVIIDNK GQRVEITNLA SKVQIKDVNS KLLLNKQNIA

!!AA_SEQUENCE 1.0

```
ID   W22410 standard; Protein; 140 AA.
AC   W22410;
DT   08-DEC-1997  (first entry)
DE   Alpha-4 integrin mouse MAb 21.6 VH region.
KW   Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6;
KW   asthma; atherosclerosis; AIDS; dementia; diabetes; tumour;
KW   metastasis; inflammatory bowel disease; rheumatoid arthritis;
KW   transplant rejection; graft versus host disease; nephritis;
KW   atopic dermatitis; psoriasis; myocardial ischaemia;
KW   acute leukocyte mediated lung injury; therapy.
OS   Mus musculus.
FH   Key             Location/Qualifiers
FT   peptide         1. .19
FT                   /label= Leader
FT   region          20. .49
FT                   /label= FR1
FT                   /note= "framework region 1"
FT   region          50. .54
FT                   /label= CDR1
FT                   /note= "complementarity determining region 1"
FT   region          55. .68
FT                   /label= FR2
FT                   /note= "framework region 2"
FT   region          69. .85
FT                   /label= CDR2
FT                   /note= "complementarity determining region 2"
FT   region          86. .117
FT                   /label= FR3
FT                   /note= "framework region 3"
FT   region          118. .131
FT                   /label= CDR3
FT                   /note= "complementarity determining region 3"
FT   region          132. .140
FT                   /label= FR4
FT                   /note= "framework region 4"
PN   WO9718838-A1.
PD   29-MAY-1997.
PF   21-NOV-1996; U18807.
PR   21-NOV-1995; US-561521.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones ST,  Leger OJ,  Saldanha J,  Yednock TA;
DR   WPI; 97-297879/27.
DR   N-PSDB; T74760.
PT   Uses of humanised alpha-4 integrin antibody - for treatment of
PT   asthma, atherosclerosis, AIDS, dementia, etc.
PS   Claim 18; Page 69-70; 107pp; English.
CC   This polypeptide comprises the heavy chain variable region (VH) of
CC   mouse anti-alpha-4 integrin monoclonal antibody 21.6.  The
CC   complementarity determining regions (CDRs) of the 21.6 VH can be
CC   incorporated into a human 21/28'CL framework to produce a claimed
CC   humanised 21.6 VH (see W22413) and a claimed humanised 21.6
CC   antibody that is used in the manufacture of a medicament for
CC   treating a disease selected from asthma, atherosclerosis, AIDS,
CC   dementia, diabetes, inflammatory bowel disease, rheumatoid
CC   arthritis, transplant rejection, graft versus host disease, tumour
CC   metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
CC   ischaemia, and acute leukocyte mediated lung injury.  The antibody
CC   may also be used in the affinity purification of alpha-4 integrin
CC   for use as a vaccine or an immunogen.  It is also useful for
CC   generating idiotypic antibodies.  The humanised antibodies of the
CC   invention have a half-life in the human circulation essentially
CC   equivalent to that of naturally occurring human antibodies.
SQ   Sequence   140 AA;

 W22410  Length: 140  February 11, 2000 15:48  Type: P  Check: 2629  ..

       1  MKCSWVMFFL MAVVTGVNSE VQLQQSGAEL VKPGASVKLS CTASGFNIKD

      51  TYIHCVKQRP EQGLEWIGRI DPANGYTKYD PKFQGKATIT ADTSSNTAYL

     101  QLSSLTSEDT AVYFCAREGY YGNYGVYAMD YWGQGTSVTV

!!AA_SEQUENCE 1.0
```

```
ID   W22413 standard; Protein; 123 AA.
AC   W22413;
DT   08-DEC-1997  (first entry)
DE   Humanised alpha-4 integrin antibody 21.6 VH Ha.
KW   Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6;
KW   asthma; atherosclerosis; AIDS; dementia; diabetes; tumour;
KW   metastasis; inflammatory bowel disease; rheumatoid arthritis;
KW   transplant rejection; graft versus host disease; nephritis;
KW   atopic dermatitis; psoriasis; myocardial ischaemia;
KW   acute leukocyte mediated lung injury; therapy.
OS   Chimeric Mus musculus;
OS   Chimeric Homo sapiens;
OS   Chimeric synthetic.
FH   Key             Location/Qualifiers
FT   region          1. .30
FT                   /label= FR1
FT                   /note= "21/28'CL framework region 1"
FT   misc_difference 27. .30
FT                   /note= 21/28'CL residues 27-30 are replaced by
FT                   those of MAb 21.6, involved in antigen
FT                   binding"
FT   region          31. .35
FT                   /label= CDR1
FT                   /note= "21.6 complementarity determining region 1"
FT   region          36. .49
FT                   /label= FR2
FT                   /note= "21/28'CL framework region 2"
FT   region          50. .66
FT                   /label= CDR2
FT                   /note= "21.6 complementarity determining region 2"
FT   region          67. .98
FT                   /label= FR3
FT                   /note= "21/28'CL framework region 3"
FT   misc_difference 72
FT                   /note= "21/28'CL Arg-72 is substd. by Ala of mouse
FT                   21.6 VL, important in supporting the CDR2
FT                   loop"
FT   region          99. .112
FT                   /label= CDR3
FT                   /note= "21.6 complementarity determining region 3"
FT   region          113. .123
FT                   /label= FR4
FT                   /note= "21/28'CL framework region 4"
PN   WO9718838-A1.
PD   29-MAY-1997.
PF   21-NOV-1996; U18807.
PR   21-NOV-1995; US-561521.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones ST,  Leger OJ,  Saldanha J,  Yednock TA;
DR   WPI; 97-297879/27.
PT   Uses of humanised alpha-4 integrin antibody - for treatment of
PT   asthma, atherosclerosis, AIDS, dementia, etc.
PS   Claim 25; Fig 7; 107pp; English.
CC   This polypeptide, designated Ha, comprises the heavy chain variable
CC   region (VH) of a humanised alpha-4 integrin antibody 21.6.  It is
CC   composed of complementarity determining regions (CDRs) from the VH
CC   region (see W22410) of mouse alpha-4 integrin monoclonal antibody
CC   21.6 and a modified human 21/28'CL framework.  It can be expressed
CC   in mammalian host cells following PCR amplification and mutagenesis
CC   of appropriate fragments of mouse and human DNA sequences.  The
CC   humanised 21.6 VH and a humanised 21.6 VL (see W22412) can be used
CC   to produce a claimed humanised 21.6 antibody that is useful in the
CC   manufacture of a medicament for treating asthma, atherosclerosis,
CC   AIDS, dementia, diabetes, inflammatory bowel disease, rheumatoid
CC   arthritis, transplant rejection, graft versus host disease, tumour
CC   metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
CC   ischaemia, and acute leukocyte mediated lung injury.  The antibody
CC   may also be used in the affinity purification of alpha-4 integrin
CC   for use as a vaccine or an immunogen.  It is also useful for
CC   generating idiotypic antibodies.  The humanised antibody has a
CC   half-life in the human circulation essentially equivalent to that
CC   of naturally occurring human antibodies.
SQ   Sequence   123 AA;
```

```
W22413  Length: 123  February 11, 2000 15:48  Type: P  Check: 9557  ..

      1  QVQLVQSGAE VKKPGASVKV SCKASGFNIK DTYIHWVRQA PGQRLEWMGR

     51  IDPANGYTKY DPKFQGRVTI TADTSASTAY MELSSLRSED TAVYYCAREG

    101  YYGNYGVYAN DYWGQGTLVT VSS

!!AA_SEQUENCE 1.0
ID   W22428 standard; Protein; 142 AA.
AC   W22428;
DT   09-DEC-1997  (first entry)
DE   Humanised alpha-4 integrin antibody 21.6 VL version Ha.
KW   Alpha-4 integrin; humanised antibody; monoclonal antibody 21.6;
KW   asthma; atherosclerosis; AIDS; dementia; diabetes; tumour;
KW   metastasis; inflammatory bowel disease; rheumatoid arthritis;
KW   transplant rejection; graft versus host disease; nephritis;
KW   atopic dermatitis; psoriasis; myocardial ischaemia;
KW   acute leukocyte mediated lung injury; therapy.
OS   Chimeric Mus musculus;
OS   Chimeric Homo sapiens;
OS   Chimeric synthetic.
FH   Key             Location/Qualifiers
FT   peptide         1. .19
FT                   /label= Leader
FT   protein         20. .142
FT                   /label= Mat_protein
FT                   /note= "VH version Ha (Claim 25)"
FT   region          20. .49
FT                   /label= FR1
FT                   /note= "21/28'CL framework region 1"
FT   region          50. .55
FT                   /label= CDR1
FT                   /note= "21.6 complementarity determining region 1"
FT   region          55. .67
FT                   /label= FR2
FT                   /note= "21/28'CL framework region 2"
FT   region          68. .85
FT                   /label= CDR2
FT                   /note= "21.6 complementarity determining region 2"
FT   region          86. .117
FT                   /label= FR3
FT                   /note= "21/28'CL framework region 3"
FT   region          118. .131
FT                   /label= CDR3
FT                   /note= "21.6 complementarity determining region 3"
FT   region          132. .142
FT                   /label= FR4
FT                   /note= "21/28'CL framework region 4"
PN   WO9718838-A1.
PD   29-MAY-1997.
PF   21-NOV-1996; U18807.
PR   21-NOV-1995; US-561521.
PA   (ATHE-) ATHENA NEUROSCIENCES INC.
PI   Bendig MM,  Jones ST,  Leger OJ,  Saldanha J,  Yednock TA;
DR   WPI; 97-297879/27.
DR   N-PSDB; T74789.
PT   Uses of humanised alpha-4 integrin antibody - for treatment of
PT   asthma, atherosclerosis, AIDS, dementia, etc.
PS   Example 6; Fig 11; 107pp; English.
CC   This polypeptide, designated Ha, comprises the heavy chain variable
CC   region (VH) of a humanised alpha-4 integrin antibody 21.6 (see also
CC   W22413).  It is composed of complementarity determining regions from
CC   the VH region (see W22410) of mouse alpha-4 integrin monoclonal
CC   antibody 21.6 and a modified human 21/28'CL framework.  It can be
CC   expressed in mammalian host cells following PCR amplification and
CC   mutagenesis of appropriate mouse and human DNA sequences.  The
CC   humanised 21.6 VH and a humanised 21.6 VL (see W22412) can be used
CC   to produce a claimed humanised 21.6 antibody that is useful in the
CC   manufacture of a medicament for treating asthma, atherosclerosis,
CC   AIDS, dementia, diabetes, inflammatory bowel disease, rheumatoid
CC   arthritis, transplant rejection, graft versus host disease, tumour
CC   metastasis, nephritis, atopic dermatitis, psoriasis, myocardial
CC   ischaemia, and acute leukocyte mediated lung injury.  The humanised
CC   antibody has a half-life in the human circulation essentially
CC   equivalent to that of naturally occurring human antibodies.
SQ   Sequence   142 AA;

 W22428  Length: 142  February 11, 2000 15:48  Type: P  Check: 9019  ..

       1  MDWTWRVFCL LAVAPGAHSQ VQLVQSGAEV KKPGASVKVS CKASGFNIKD

      51  TYIHWVRQAP GQRLEWMGRI DPANGYTKYD PKFQGRVTIT ADTSASTAYM

     101  ELSSLRSEDT AVYYCAREGY YGNYGVYAMD YWGQGTLVTV SS

!!AA_SEQUENCE 1.0
ID   W26413 standard; Protein; 121 AA.
AC   W26413;
DT   16-DEC-1997  (first entry)
DE   Swinepox virus HindIII C encoded protein C22R.
KW   SPV; vaccine; vector; pseudorabies virus; C22R.
OS   Swinepox virus.
PN   US5651972-A.
PD   29-JUL-1997.
PF   21-APR-1989; 342212.
PR   01-JUL-1992; US-908241.
PR   21-APR-1989; US-342212.
PR   29-JUN-1992; US-908630.
PR   14-SEP-1994; US-307499.
PA   (UYFL ) UNIV FLORIDA RES FOUND INC.
PI   Gibbs EPJ,  Moyer RW,  Vinuela E;
DR   WPI; 97-392897/36.
DR   N-PSDB; T84564.
PT   Recombinant swinepox virus vector - used particularly for vaccines
PT   against infectious agents, including pseudorabies
PS   Example 2; Column 45-48; 70pp; English.
CC   This sequence comprises a polypeptide, designated C22R, encoded
CC   by the sense strand (open reading frame 3) of the HindIII C
CC   fragment (see T84564) of swinepox virus (SPV).  C22R shows no
CC   apparent homology to known protein sequences.  A claimed
CC   recombinant vector comprises a heterologous nucleotide sequence
CC   inserted into, or replacing, all or a portion of a non-essential
CC   SPV gene or nucleic acid sequence of the HindIII C fragment.  The
CC   vector can be used for the expression of heterologous proteins,
CC   both in vivo as a vaccine, and in vitro for production of the
CC   selected protein.  The heterologous protein is preferably
CC   pseudorabies virus gp50 or gp63 for use in swine vaccines.  As SPV
CC   is host-restricted to swine, the use of modified recombinant SPV as
CC   a live vaccine vector eliminates the risk of spreading infection
CC   with the virus to other animal populations.
SQ   Sequence   121 AA;

 W26413  Length: 121  February 11, 2000 15:48  Type: P  Check: 3737  ..

       1  MIIDVITSLD GGCNKCLNIK HAKSQCKPSS LLISSLENVN PGINPRFFIQ

      51  NIAANEPEKN IPSTAANATT LSPYTLLLDI HFLAQSAFFF THSIVSIALN

     101  KCIFSILSLT YVSINKLYIS E

!!AA_SEQUENCE 1.0
ID   W17971 standard; Protein; 47 AA.
AC   W17971;
DT   12-DEC-1997  (first entry)
DE   RAC protein kinase C-terminal binding protein C-terminal region.
KW   RAC protein kinase C-terminal binding protein; CTBP; modulator;
KW   signal transduction; insulin; cell proliferation; glycogen.
OS   Homo sapiens.
PN   WO9718303-A1.
PD   22-MAY-1997.
PF   05-NOV-1996; E04810.
PR   15-DEC-1995; GB-025704.
PR   16-NOV-1995; GB-023379.
PA   (NOVS ) NOVARTIS AG.
```

```
PI   Hemmings BA;
DR   WPI; 97-289279/26.
DR   N-PSDB; T67134.
PT   RAC protein kinase, or modulator excluding wortmannin and vanadate,
PT   for use as medicament - and screening potential modulators of
PT   insulin mediated intracellular signalling using RAC-PK, or fragment
PS   Claim 8; Page 25; 38pp; English.
CC   This polypeptide comprises the putative C-terminal region of a
CC   novel RAC protein kinase C-terminal binding protein (CTBP).  Its
CC   sequence was deduced from a cDNA clone (T67134) isolated from a
CC   HeLa library following a yeast two-hybrid screen assay for clones
CC   that showed specific interaction with RAC's kinase domain with its
CC   C-terminal extension.  The C-terminal domain of RAC protein kinase
CC   is phosphorylated in response to insulin activation, suggesting a
CC   role for CTBP as a modulator of insulin action.  RAC protein kinase
CC   (see W17972) and modulators of insulin-mediated intracellular
CC   signalling such as CTBP can be used in the treatment of
CC   abnormalities of cellular metabolism, diseases involving an anomaly
CC   in insulin response, and diseases involving an anomaly in glycogen
CC   metabolism.
SQ   Sequence   47 AA;

 W17971  Length: 47  February 11, 2000 15:48  Type: P  Check: 6033  ..

       1  NSARARASAA PRPGAMLRSC AARLRTLGAL CRPPVGRRLP EATRDPS

!!AA_SEQUENCE 1.0
ID   W34668 standard; Protein; 438 AA.
AC   W34668;
DT   13-FEB-1998  (first entry)
DE   Arabidopsis thaliana Rec-A protein.
KW   Rec-A protein; Synechococcus recA gene; transit peptide;
KW   homologous recombination; DNA repair; the recA gene.
OS   Arabidopsis thaliana.
FH   Key             Location/Qualifiers
FT   Peptide         1. .50
FT                   /label= signal_peptide
FT                   /note= "putative chloroplast transit peptide"
FT   Protein         51. .438
FT                   /label= mature_peptide
PN   US5674992-A.
PD   07-OCT-1997.
PF   16-JUN-1994; 938332.
PR   28-AUG-1992; US-938332.
PA   (CORR ) CORNELL RES FOUND INC.
PI   Cerutti H,  Jagendorf A;
DR   WPI; 97-502391/46.
DR   N-PSDB; T93777.
PT   DNA encoding chloroplast RecA-like protein from Arabidopsis thaliana
PT   - useful for increasing the frequency of homologous chromosomal
PT   recombination when lacking the transit peptide sequence
PS   Claim 2; Columns 7-9; 14pp; English.
CC   The present sequence is of a Rec-A like protein from Arabidopsis
CC   thaliana. The cDNA encoding this protein (T93777) was isolated from an
CC   Arabidopsis library screened with a Synechococcus recA gene. The cDNA
CC   sequence is truncated at its 5' end, and the start of the open reading
CC   frame, and thus the start of the transit peptide, is missing. In E. coli,
CC   and many other prokaryotes, the Rec-A protein is essential for homologous
CC   recombination, and DNA repair. In view of this role, it is likely that
CC   that the enzyme encoded by the present sequence may also be concerned
CC   with DNA repair. The 5' end of the recA gene, i.e. the portion encoding
CC   the transit peptide part of the protein that directs the final product
CC   into chloroplast, could be removed and replaced by an amino acid sequence
CC   that directs completed proteins to enter the nucleus. By introducing the
CC   RecA protein into the nuclei of higher plants, it should be possible to
CC   increase the frequency of homologous recombination in the chromosomal
CC   genes of the plants.
SQ   Sequence   438 AA;

W34668  Length: 438  February 11, 2000 15:48  Type: P  Check: 9097  ..

       1  DSQLVLSLKL NPSFTPLSPL FPFTPCSSFS PSLRFSSCYS RRLYSPVTVY

      51  AAKKLSHKIS SEFDDRINGA LSPDADSRFL DRQKALEAAM NDINSSFGKG

     101  SVTRLGSAGG ALVETFSSGI LTLDLALGGG LPKGRVVEIY GPESSGKTTL

     151  ALHAIAEVQK LGGNAMLVDA EHAFDPAYSK ALGVDVENLI VCQPDNGEMA

     201  LETADRMCRS GAVDLICVDS VSALTPRAEI EGEIGMQQMG LQARLMSQAL

     251  RKMSGNASKA GCTLIFLNQI RYKIGVYYGN PEVTSGGIAL KFFASVRLEI

     301  RSAGKIKSSK GDEDIGLRAR VRVQKSKVSR PYKQAEFEIM FGEGVSKLGC

     351  VLDCAEIMEV VVKKGSWYSY EDQRLGQGRE KALQHLRENP ALQDEIEKKV

     401  RLLMLDGEVH RSTPLMSSSS SSASHREEEE EDSLDDFQ

!!AA_SEQUENCE 1.0
ID   W24491 standard; peptide; 122 AA.
AC   W24491;
DT   27-FEB-1998  (first entry)
DE   Novel amylase inhibitor protein subunit isolated from wheat.
KW   Amylase inhibitor; wheat flour; wheat; human alpha-amylase;
KW   blood sugar level; insulin secretion; fatty acid level; appetite
KW   suppressor; treatment; diabetes; prevention; obesity.
OS   Triticum sp.
PN   EP-785214-A2.
PD   23-JUL-1997.
PF   10-JAN-1997; 100316.
PR   18-JAN-1996; JP-023446.
PA   (NAGA-) NAGATA SANGYO CO LTD.
PA   (NISS ) NISSHIN FLOUR MILLING CO.
PI   Miyazaki T,  Morimoto T,  Murayama R;
DR   WPI; 97-365900/34.
PT   New isolated proteins from wheat active as amylase inhibitors - used
PT   in the treatment and prevention of diabetes and in the prevention of
PT   obesity
PS   Claim 1; Page 16; 22pp; English.
CC   This sequence represents a novel amylase inhibitory protein which is
CC   composed of two subunits, each having the present sequence. This protein
CC   can be prepared by extraction from wheat flour, wheat or wheat gluten.
CC   The protein is spherical and has S-S bonds, and has a high inhibitory
CC   activity against human alpha-amylase. It has about 5-30 times higher
CC   activity as compared with several known amylase inhibitors. It can also
CC   inhibit increases in blood sugar levels, inhibit insulin secretion and
CC   inhibit an increase in fatty acid levels in the blood, thus maintaining a
CC   feeling of satiety after meals, and easily suppressing appetite. The
CC   proteins can be used in the treatment and prevention of diabetes and the
CC   prevention of obesity.
SQ   Sequence   122 AA;

 W24491  Length: 122  February 11, 2000 15:48  Type: P  Check: 1423  ..

       1  SGPWMCYPGY AFKVPALPGC RPVLKLQCNG SQVPEAVLRD CCQQLADISE

      51  WCRCGALYSM LDSMYKEHGV QEGQAGTGAF PSCRREVVKL TAASITAVCK

     101  LPIVIDASGD GAYVCKGVAA YP

!!AA_SEQUENCE 1.0
ID   W24492 standard; peptide; 124 AA.
AC   W24492;
DT   27-FEB-1998  (first entry)
DE   Novel amylase inhibitor protein 2 subunit isolated from wheat.
KW   Amylase inhibitor; wheat flour; wheat; human alpha-amylase;
KW   blood sugar level; insulin secretion; fatty acid level; appetite
KW   suppressor; treatment; diabetes; prevention; obesity.
OS   Triticum sp.
PN   EP-785214-A2.
PD   23-JUL-1997.
PF   10-JAN-1997; 100316.
PR   18-JAN-1996; JP-023446.
PA   (NAGA-) NAGATA SANGYO CO LTD.
PA   (NISS ) NISSHIN FLOUR MILLING CO.
```

PI Miyazaki T, Morimoto T, Murayama R;
DR WPI; 97-365900/34.
PT New isolated proteins from wheat active as amylase inhibitors - used
PT in the treatment and prevention of diabetes and in the prevention of
PT obesity
PS Claim 5; Page 17; 22pp; English.
CC This sequence represents a novel amylase inhibitory protein which is
CC composed of two subunits, each having the present sequence. This protein
CC can be prepared by extraction from wheat flour, wheat or wheat gluten.
CC The protein is spherical and has S-S bonds, and has a high inhibitory
CC activity against human alpha-amylase. It has a higher inhibitory
CC activity as compared with several known amylase inhibitors. It can also
CC inhibit increases in blood sugar levels, inhibit insulin secretion and
CC inhibit an increase in fatty acid levels in the blood, thus maintaining a
CC feeling of satiety after meals, and easily suppressing appetite. The
CC proteins can be used in the treatment and prevention of diabetes and the
CC prevention of obesity.
SQ Sequence 124 AA;

W24492 Length: 124 February 11, 2000 15:48 Type: P Check: 2685 ..

1 SGPWMCYPGY AFKVPALPGC RPVLKLQCNG SQVPEAVLRD CCQQLADISE

51 WCRCGALYSM LDSMYKEHGV QEGQAGTGAF PSCRREVVKL TAASITAVCK

101 LPIVIDASGD GAYVCKGVAA YPDA

!!AA_SEQUENCE 1.0
ID W24493 standard; peptide; 124 AA.
AC W24493;
DT 27-FEB-1998 (first entry)
DE Novel amylase inhibitor protein 3 subunit isolated from wheat.
KW Amylase inhibitor; wheat flour; wheat; human alpha-amylase;
KW blood sugar level; insulin secretion; fatty acid level; appetite
KW suppressor; treatment; diabetes; prevention; obesity.
OS Triticum sp.
PN EP-785214-A2.
PD 23-JUL-1997.
PF 10-JAN-1997; 100316.
PR 18-JAN-1996; JP-023446.
PA (NAGA-) NAGATA SANGYO CO LTD.
PA (NISS) NISSHIN FLOUR MILLING CO.
PI Miyazaki T, Morimoto T, Murayama R;
DR WPI; 97-365900/34.
PT New isolated proteins from wheat active as amylase inhibitors - used
PT in the treatment and prevention of diabetes and in the prevention of
PT obesity
PS Claim 5; Page 18; 22pp; English.
CC This sequence represents a novel amylase inhibitory protein which is
CC composed of two subunits, each having the present sequence. This protein
CC can be prepared by extraction from wheat flour, wheat or wheat gluten.
CC The protein is spherical and has S-S bonds, and has a high inhibitory
CC activity against human alpha-amylase. It has about a higher inhibitory
CC activity as compared with several known amylase inhibitors. It can also
CC inhibit increases in blood sugar levels, inhibit insulin secretion and
CC inhibit an increase in fatty acid levels in the blood, thus maintaining a
CC feeling of satiety after meals, and easily suppressing appetite. The
CC proteins can be used in the treatment and prevention of diabetes and the
CC prevention of obesity.
SQ Sequence 124 AA;

W24493 Length: 124 February 11, 2000 15:48 Type: P Check: 3320 ..

1 SGPWMCYPGQ AFQVPALPAC RPLLRLQCNG SQVPEAVLRD CCQQLAHISE

51 WCRCGALYSM LDSMYKEHGA QEGQAGTGAF PRCRREVVKL TAASITAVCR

101 LPIVVDASGD GAYVCKDVAA YPDA

!!AA_SEQUENCE 1.0
ID W29523 standard; Protein; 124 AA.
AC W29523;
DT 09-MAR-1998 (first entry)
DE Wheat amylase inhibitor 0.26 Ala subunit.
KW Amylase inhibitor; 0.26 Ala; wheat; visceral fat; obesity; therapy.
OS Triticum aestivum.
PN EP-784978-A2.
PD 23-JUL-1997.
PF 11-JAN-1997; 100399.
PR 18-JAN-1996; JP-023445.
PA (NAGA-) NAGATA SANGYO CO LTD.
PA (NISS) NISSHIN FLOUR MILLING CO.
PI Goda T, Miyazaki T, Morimoto T, Murayama R, Takase S;
DR WPI; 97-365735/34.
PT Use of an amylase inhibitor from wheat - for inhibiting the
PT accumulation of visceral fat to prevent visceral fat obesity
PS Claim 2; Page 16; 22pp; English.
CC This protein comprises the subunit of the wheat amylase
CC inhibitor homodimer protein 0.26 Ala. An agent for inhibiting an
CC accumulation of visceral fat and hence for preventing visceral
CC fat obesity is claimed. It contains an amylase inhibitor of wheat
CC origin as an active ingredient. The amylase inhibitor is selected
CC from 0.26 Ala, 0.26 Alb (see W29524) and 0.19 Al (see W29525). The
CC 0.26 Ala protein has inhibitory activity of 26,100 U/mg against
CC human pancreatic alpha-amylase. Also claimed is a food additive
CC comprising 0.26 Alb. The amylase inhibitors can inhibit the
CC digestion of starch, inhibit increases in blood glucose level,
CC reduce insulin secretion and reduce the activity of lipogenic
CC enzymes in the viscus.
SQ Sequence 124 AA;

W29523 Length: 124 February 11, 2000 15:48 Type: P Check: 2685 ..

1 SGPWMCYPGY AFKVPALPGC RPVLKLQCNG SQVPEAVLRD CCQQLADISE

51 WCRCGALYSM LDSMYKEHGV QEGQAGTGAF PSCRREVVKL TAASITAVCK

101 LPIVIDASGD GAYVCKGVAA YPDA

!!AA_SEQUENCE 1.0
ID W29524 standard; Protein; 122 AA.
AC W29524;
DT 09-MAR-1998 (first entry)
DE Wheat amylase inhibitor 0.26 Alb subunit.
KW Amylase inhibitor; 0.26 Alb; wheat; visceral fat; obesity; therapy.
OS Triticum aestivum.
PN EP-784978-A2.
PD 23-JUL-1997.
PF 11-JAN-1997; 100399.
PR 18-JAN-1996; JP-023445.
PA (NAGA-) NAGATA SANGYO CO LTD.
PA (NISS) NISSHIN FLOUR MILLING CO.
PI Goda T, Miyazaki T, Morimoto T, Murayama R, Takase S;
DR WPI; 97-365735/34.
PT Use of an amylase inhibitor from wheat - for inhibiting the
PT accumulation of visceral fat to prevent visceral fat obesity
PS Claim 2; Page 17; 22pp; English.
CC This protein comprises the subunit of the novel wheat amylase
CC inhibitor homodimer protein 0.26 Alb. An agent for inhibiting an
CC accumulation of visceral fat and hence for preventing visceral
CC fat obesity is claimed. It contains an amylase inhibitor of wheat
CC origin as an active ingredient. The amylase inhibitor is selected
CC from 0.26 Alb, 0.26 Ala (see W29523) and 0.19 Al (see W29525). The
CC 0.26 Alb protein has inhibitory activity of 20,500 U/mg against
CC human pancreatic alpha-amylase. Also claimed is a food additive
CC comprising 0.26 Alb. The amylase inhibitors can inhibit the
CC digestion of starch, inhibit increases in blood glucose level,
CC reduce insulin secretion and reduce the activity of lipogenic
CC enzymes in the viscus.
SQ Sequence 122 AA;

W29524 Length: 122 February 11, 2000 15:48 Type: P Check: 1423 ..

1 SGPWMCYPGY AFKVPALPGC RPVLKLQCNG SQVPEAVLRD CCQQLADISE

51 WCRCGALYSM LDSMYKEHGV QEGQAGTGAF PSCRREVVKL TAASITAVCK

```
     101  LPIVIDASGD GAYVCKGVAA YP

!!AA_SEQUENCE 1.0
ID   W29525 standard; Protein; 124 AA.
AC   W29525;
DT   09-MAR-1998  (first entry)
DE   Wheat amylase inhibitor 0.19 Al subunit.
KW   Amylase inhibitor; 0.19 Al; wheat; visceral fat; obesity; therapy.
OS   Triticum aestivum.
PN   EP-784978-A2.
PD   23-JUL-1997.
PF   11-JAN-1997; 100399.
PR   18-JAN-1996; JP-023445.
PA   (NAGA-) NAGATA SANGYO CO LTD.
PA   (NISS ) NISSHIN FLOUR MILLING CO.
PI   Goda T,  Miyazaki T,  Morimoto T,  Murayama R,  Takase S;
DR   WPI; 97-365735/34.
PT   Use of an amylase inhibitor from wheat - for inhibiting the
PT   accumulation of visceral fat to prevent visceral fat obesity
PS   Claim 2; Page 18; 22pp; English.
CC   This protein comprises the subunit of the wheat amylase
CC   inhibitor homodimer protein 0.19 Al.  An agent for inhibiting an
CC   accumulation of visceral fat and hence for preventing visceral
CC   fat obesity is claimed.  It contains an amylase inhibitor of wheat
CC   origin as an active ingredient.  The amylase inhibitor is selected
CC   from 0.26 Ala (see W29523), 0.26 Alb (see W29524) and 0.19 Al.  The
CC   0.19 Al protein has inhibitory activity of 20,300 U/mg against
CC   human pancreatic alpha-amylase.  Also claimed is a food additive
CC   comprising 0.26 Alb.  The amylase inhibitors can inhibit the
CC   digestion of starch, inhibit increases in blood glucose level,
CC   reduce insulin secretion and reduce the activity of lipogenic
CC   enzymes in the viscus.
SQ   Sequence   124 AA;

 W29525  Length: 124  February 11, 2000 15:48  Type: P  Check: 3320  ..

       1  SGPWMCYPGQ AFQVPALPAC RPLLRLQCNG SQVPEAVLRD CCQQLAHISE

      51  WCRCGALYSM LDSMYKEHGA QEGQAGTGAF PRCRREVVKL TAASITAVCR

     101  LPIVVDASGD GAYVCKDVAA YPDA

!!AA_SEQUENCE 1.0
ID   R63906 standard; Protein; 248 AA.
AC   R63906;
DT   27-JUL-1995  (first entry)
DE   Type I ribosome-inactivating protein luffin.
KW   Type I ribosome-inactivating proteins; luffin; RIPs;
KW   cytotoxic therapeutic agents; autoimmune disease; cancer;
KW   graft-versus-host disease.
OS   Luffa sp.
PN   WO9426910-A.
PD   24-NOV-1994.
PF   12-MAY-1994; U05348.
PR   12-MAY-1993; US-064691.
PA   (XOMA ) XOMA CORP.
PI   Better MD,  Carroll SS,  Studnicka GM,  Carroll SF;
DR   WPI; 95-006804/01.
PT   Polynucleotide(s) encoding type I ribosome-inactivating proteins
PT   - which are suitable for use as components of cytotoxic
PT   therapeutic agents.
PS   Disclosure; Fig 4; 221pp; English.
CC   R63906 is the type I ribosome-inactivating protein (RIP)
CC   luffin, one of the nine RIPs described in R63903-R63911.
CC   RIPs are key components of cytotoxic therapeutic agents (CTAs),
CC   which include gene fusion products and immunoconjugates. CTAs may
CC   be used to selectively eliminate any cell type to which a RIP
CC   component is targetted, by the specific binding capacity of the
CC   second component of the agent. They can be used in the treatment
CC   of diseases where the elimination of a particular cell type is
CC   desired, such as autoimmune disease, cancer and graft-versus-host
CC   disease.
SQ   Sequence   248 AA;

 R63906  Length: 248  February 11, 2000 15:48  Type: P  Check: 7673  ..

       1  DVRFSLSGSS STSYSKFIGD LRKALPSNGT VYNLTILLSS ASGASRYTLM

      51  TLSNYDGKAI TVAVDVSQLY IMGYLVNSTS YFFNESDAKL ASQYVFKGST

     101  IVTLPYSGNY EKLQTAAGKI REKIPLGFPA LDSALTTIFH YDSTAAAAAF

     151  LVILQTTAEA SRFKYIEGQI IERISKNQVP SLATISLENS LWSALSKQIQ

     201  LAQTNNGTFK TPVVITDDKG QRVEITNVTS KVVTKNIQLL LNYKQNVA

!!AA_SEQUENCE 1.0
ID   W45093 standard; peptide; 176 AA.
AC   W45093;
DT   07-MAY-1998  (first entry)
DE   Residues 217-392 of human type B EBNA2 (strain AG876).
KW   CBF1/RBPJk interaction domain; EBNA2; transplant.
OS   Epstein-Barr virus.
FH   Key             Location/Qualifiers
FT   Region          69. .76
FT                   /label= conserved_region_6
PN   WO9730081-A1.
PD   21-AUG-1997.
PF   12-FEB-1997; U02243.
PR   13-FEB-1996; US-011557.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Hayward SD, Ling PD;
DR   WPI; 97-424974/39.
PT   Peptide with homology to CBF1-RBPJk interaction domain of EBNA2 -
PT   useful to provide protection against Epstein-Barr virus
PS   Disclosure; Fig 5; 62pp; English.
CC   The patent discloses five new peptides which display substantial
CC   homology with the CBF1/RBPJk interaction domain of EBNA2 (see
CC   W24280-W24284). In particular, the peptides include amino acids
CC   318-327 of EBNA2 (i.e. conserved region 6, or CR6). The peptides,
CC   when administered in an amount sufficient to inhibit interaction
CC   of EBV EBNA2 with CBF1, are used to protect mammals prior to or
CC   following a transplant, against Epstein-Barr virus. Labelled
CC   peptides are used to detect factors that interact with EBV EBNA2.
CC   The present sequence is from one strain of EB virus (human type B)
CC   used to identify CR6 (GPPWWPPV; residues 69-76).
SQ   Sequence   176 AA;

 W45093  Length: 176  February 11, 2000 15:48  Type: P  Check: 9381  ..

       1  HQLSLPPHPP PHQSTPHCSS DSTGLPPPPT SYSIPSMTLS PEPLPPPAAP

      51  AHPLPGVIYD QQALPPTPGP PWWPPVRDPT PTTQTPPTNT KQPGDQGQGR

     101  GRWRGRGRSK GRGRMHKLPE PRRPGPDTSS PSMPQLSPVV SLHQGQGPEN

     151  SPTPGPSTAG PVCRVTPSAT PDISPI

!!AA_SEQUENCE 1.0
ID   W38190 standard; Protein; 107 AA.
AC   W38190;
DT   11-MAY-1998  (first entry)
DE   Soybean SCARECROW SRPg1 protein.
KW   SCARECROW; SCR gene; SRPg1; transgenic plant; root; gravitropism;
KW   crop improvement; soybean.
OS   Glycine max.
FH   Key             Location/Qualifiers
FT   Misc_difference 23
FT                   /note= "encoded by NNC"
PN   WO9741152-A1.
PD   06-NOV-1997.
PF   25-APR-1997; U07022.
PR   24-APR-1997; US-842445.
PR   26-APR-1996; US-638617.
PA   (UYNY ) UNIV NEW YORK STATE.
```

PI Benfey PN, Dilaurenzio L, Helariutta Y, Malamy JE,
PI Pysh L, Wysocka-Diller J;
DR WPI; 97-549683/50.
DR N-PSDB; T95765.
PT DNA encoding Arabidopsis SCARECROW protein - useful to modify plant
PT cell division and therefore alter root development, or alter plant
PT stem or hypocotyl gravitropism
PS Claim 2; Page 136; 221pp; English.
CC This is the deduced amino acid sequence of soybean genomic clone
CC SRPg1 (see T95765). It shows homology to Arabidopsis SCARECROW
CC (SCR) protein (see W38178). SCR is a member of a novel protein
CC family and contains a number of potential functional domains
CC similar to those found in transcription factors. SCR is expressed
CC specifically in embryo root progenitor tissue and in certain root
CC and stem tissues. It regulates a specific asymmetric division, and
CC controls gravitropic response in aerial structures and root
CC formation. SCR proteins (see W38178-201) from dicots and monocots,
CC SCR proteins lacking 1-4 of MOTIFs I-VI, and polypeptides
CC corresponding to MOTIF I, II, III, IV, V or VI of SCR are claimed.
CC Transgenic plants can be engineered to overexpress the SCR protein,
CC so that cell division is increased in roots, resulting in thicker
CC root development, while a plant with an altered stem or hypocotyl
CC gravitropism is less susceptible to lodging. Plants that contain
CC an antisense molecule that suppresses the expression of endogenous
CC SCR gene product have thinner root development. A gene of interest
CC can be placed under control of a SCR promoter and expressed in a
CC plant to confer herbicide, salt, pathogen or insect resistance, or
CC when expressed in stems to increase starch, lignin or cellulose
CC biosynthesis (all claimed).
SQ Sequence 107 AA;

W38190 Length: 107 February 11, 2000 15:48 Type: P Check: 1840 ..

1 STENLEDANK MLLEISQLST PFXTSAQRVA AYFSEAISAR LVSSCLGIYA

51 TLPHTHQSHK VASAFQVFNG ISPLVEFSHF TANQAIQEAF EREERVHIID

101 LDIMQGL

!!AA_SEQUENCE 1.0
ID W34214 standard; Protein; 271 AA.
AC W34214;
DT 11-MAY-1998 (first entry)
DE Streptomyces ketoreductase (frenolicin gene P product).
KW Frenolicin; antibiotic; feed additive; anticoccidial;
KW coccidiostatic; ketoreductase; Streptomyces roseofulvus.
OS Streptomyces sp.
PN EP-806480-A2.
PD 12-NOV-1997.
PF 02-MAY-1997; 107329.
PR 07-MAY-1996; US-016753.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PI Reeves CD, Soliday CL;
DR WPI; 97-538619/50.
DR N-PSDB; T93095.
PT Streptomyces frenolicin gene cluster - useful for producing
PT recombinant frenolicin antibiotics
PS Claim 8; Page 33-34; 66pp; English.
CC This is a ketoreductase related to the actinorhodin actIII protein.
CC It is the encoded product of gene P of a claimed Streptomyces
CC frenolicin gene cluster (see T93095) that contains specifically
CC claimed coding sequences (genes A-U) that respectively encode 21
CC proteins (see W34199-219) involved in frenolicin synthesis. The 21
CC genes of the cluster can be divided into 5 subclusters: (1) genes
CC A, B, C, D and F encode an efflux pump; (2) genes H, I, J and K
CC encode butyrate starter synthases; (3) genes L, M and N encode
CC polyketide synthases; (4) genes O, P, Q and R encode a hemiketalase,
CC a ketoreductase and cyclases/dehydrases; and (5) genes S and T
CC encode a keto/enoyl reductase and a hydrolase. Also claimed are
CC vectors and host cells (especially a Streptomyces sp., particularly
CC Streptomyces roseofulvus). Host cells can be cultured to produce
CC frenolicins or frenolicin precursors. The precursors can be
CC converted to frenolicins by chemical or other methods. The
CC frenolicins can be oxidised to frenolicin B, an antibiotic used as
CC an anticoccidial agent. The frenolicins can be used as animal feed
CC additives.
SQ Sequence 271 AA;

W34214 Length: 271 February 11, 2000 15:48 Type: P Check: 5054 ..

1 MTTAAPHTRP GEAGTTRGPA LVTGATRGIG LAVAEALVAR GYPVVVCARD

51 AEAVARTVKE LAAGGARVEG VVADVTDAAS VHELVATTVA RFGPVEVLVN

101 NAGRSGGGVT AELSESLWDD VIATNLKSVF LVTREVLTTG GMTGRGRGVV

151 NIASTGGKQG VVFGAPYSAS KHGVVGFTKA LGLELARSGI TVNAVCPGYV

201 ETPMAAGVRR HYADLWDVTE EDVLARFEAK IPLGRYTRPD EVAALVDYLV

251 TDAAAAVTAQ ALNVCGGLGN Y

!!AA_SEQUENCE 1.0
ID W41747 standard; protein; 239 AA.
AC W41747;
DT 22-MAY-1998 (first entry)
DE Hepatitis C virus antigen.
KW non-A non-B hepatitis virus; NANBH; hepatitis C virus; HCV;
KW antigen; diagnosis; detection.
OS Hepatitis virus.
PN J05176774-A.
PD 20-JUL-1993.
PF 18-DEC-1991; 354708.
PR 18-DEC-1990; JP-412020.
PA (GREC) GREEN CROSS CORP.
PA (SHIM/) SHIMOTONO K.
DR WPI; 93-260858/33.
PT Protein contg. non-A non-B hepatitis antigen fragment - prepd. by
PT culturing transformants transformed by vector contg. base sequence
PT coding specified aminoacid sequences, used for detecting hepatitis
PS Claim 1; Fig 13; 53pp; Japanese.
CC The present sequence is a non-A non-B hepatitis virus (NANBH) or
CC hepatitis C virus (HCV) antigen, useful for diagnosis or detection.
SQ Sequence 239 AA;

W41747 Length: 239 February 11, 2000 15:48 Type: P Check: 3783 ..

1 VCCSMSYTWT GALITPCAAE ESKLPINPLS NSLLRHHSMV YSTTSRSASL

51 RQKKVTFDRL QVLDDHYRDV LKEMKAKAST VKARLLSIEE ACKLTPPHSA

101 KSKFGYGAKD VRSLSSRAVN HIRSVWEDLL EDTETPIDTT IMAKNEVFCV

151 QPEKGGRKPA RLIVFPDLGV RVCEKMALYD VVSTLPQAVM GPSYGFQYSP

201 GQRVEFLVNT WKSKKCPMGF SYDTRCFDST VTENDIRTE

!!AA_SEQUENCE 1.0
ID W47606 standard; Protein; 194 AA.
AC W47606;
DT 11-JUN-1998 (first entry)
DE HRSV protein 22K.
KW HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine;
KW protein 22K.
OS Human respiratory syncytial virus.
PN US5716823-A.
PD 10-FEB-1998.
PF 12-MAY-1997; 854783.
PR 13-JUL-1988; US-218737.
PR 14-JAN-1986; US-818740.
PR 23-DEC-1986; WO-U02756.
PR 11-JUN-1992; US-897171.
PR 12-MAY-1997; US-854783.
PA (PHAA) PHARMACIA & UPJOHN CO.
PI Collins PL, Wertz GW;
DR WPI; 98-144802/13.

```
DR   N-PSDB; V18737.
PT   Production of human respiratory syncytial virus glyco-protein F or G
PT   - by culturing eukaryotic host cells transfected with corresponding
PT   DNA
PS   Example 1; Columns 27-30; 17pp; English.
CC   The present sequence was used in the development of a novel method
CC   for the production of human respiratory syncytial virus (HRSV)
CC   glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
CC   culturing eukaryotic host cells transfected with an isolated DNA
CC   sequence encoding HRSV gpF or gpG. The gp can be used to prepare
CC   vaccines against HRSV.
SQ   Sequence   194 AA;

 W47606  Length: 194  February 11, 2000 15:48  Type: P  Check: 9828  ..

       1  MSRRNPCKFE IRGHCLNGKR CHFSHNYFEW PPHALLVRQN FMLNRILKSM

      51  DKSIDTLSEI SGAAELDRTE EYALGVVGVL ESYIGSINNI TKQSACVAMS

     101  KLLTELNSDD IKKLRDNEEL NSPKIRVYNT VISYIESNRK NNKQTIHLLK

     151  RLPADVLKKT IKNTLDIHKS ITINNPKEST VSDTNDHAKN NDTT

!!AA_SEQUENCE 1.0
ID   W50006 standard; Protein; 365 AA.
AC   W50006;
DT   11-JUN-1998  (first entry)
DE   Human papillomavirus-18 E2 protein.
KW   Human papillomavirus-18; HPV-18; E2 protein; infection; treatment;
KW   wart; epidermodysplasia verruciformis; laryngeal papilloma;
KW   cervical carcinoma.
OS   Human papillomavirus.
FH   Key             Location/Qualifiers
FT   Domain          1. .203
FT                   /note= "transcription activating domain"
FT   Domain          282. .365
FT                   /note= "DNA-binding domain"
PN   WO9801148-A1.
PD   15-JAN-1998.
PF   08-JUL-1997; U11815.
PR   09-JUL-1996; US-677206.
PA   (HARD ) HARVARD COLLEGE.
PI   Dowhanick JJ, Howley PM;
DR   WPI; 98-100815/09.
DR   N-PSDB; V18818.
PT   Treating papilloma virus infections, by repressing E6 and E7
PT   oncogenic proteins - using E2 polypeptide or nucleic acid encoding
PT   it, useful for treating warts and cervical cancer
PS   Claim 22; Pages 51-53; 71pp; English.
CC   The present sequence is the human papillomavirus-18 (HPV-18)
CC   E2 protein.
CC   Papillomavirus infections can be treated by administering an E2
CC   protein, comprising a DNA-binding and transcription activating
CC   domain, or the gene encoding the E2 protein to infected cells. The
CC   method is applicable to livestock, zoo animals, pets and humans,
CC   specifically for the treatment, prevention or reversal of warts,
CC   e.g. plantar, common, Butcher's common, flat or genital warts, or
CC   epidermodysplasia verruciformis, or growth of papillomavirus
CC   transformed cells, especially laryngeal papilloma or cervical
CC   carcinoma.
SQ   Sequence   365 AA;

 W50006  Length: 365  February 11, 2000 15:48  Type: P  Check: 5567  ..

       1  MQTPKETLSE RLSCVQDKII DHYENDSKDI DSQIQYWQLI RWENAIFFAA

      51  REHGIQTLNH QVVPAYNISK SKAHKAIELQ MALQGLAQSR YKTEDWTLQD

     101  TCEELWNTEP THCFKKGGQT VQVYFDGNKD NCMTYVAWDS VYYMTDAGTW

     151  DKTATCVSHR GLYYVKEGYN TFYIEFKSEC EKYGNTGTWE VHFGNNVIDC

     201  NDSMCSTSDD TVSATQLVKQ LQHTPSPYSS TVSVGTAKTY GQTSAATRPG

     251  HCGLAEKQHC GPVNPLLGAA TPTGNNKRRK LCSGNTTPII HLKGDRNSLK

     301  CLRYRLRKHS DHYRDISSTW HWTGAGNEKT GILTVTYHSE TQRTKFLNTV

     351  AIPDSVQILV GYMTM

!!AA_SEQUENCE 1.0
ID   W55590 standard; Protein; 307 AA.
AC   W55590;
DT   07-JUL-1998  (first entry)
DE   H. pylori ORF 09cp10224_429510_c2_46.aa cell envelope OMP.
KW   Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW   identification; binding compound; bacteria; life cycle; activator;
KW   inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS   Helicobacter pylori.
PN   WO9737044-A1.
PD   09-OCT-1997.
PF   27-MAR-1997; U05223.
PR   06-DEC-1996; US-761318.
PR   29-MAR-1996; US-625811.
PR   02-APR-1996; US-758731.
PR   25-OCT-1996; US-736905.
PR   28-OCT-1996; US-738859.
PA   (ASTR ) ASTRA AB.
PI   Alm RA, Smith D;
DR   WPI; 97-503122/46.
DR   N-PSDB; V24999.
PT   Helicobacter pylori nucleic acid sequences and encoded
PT   polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT   infection and for diagnosis of H. pylori infection
PS   Claims 14,80; Pages 792-793; 1145pp; English.
CC   This sequence is a H. pylori cell envelope outer membrane protein (OMP)
CC   having a C-terminal tyrosine cluster motif. The protein may be used in a
CC   vaccine to prevent or treat H. pylori infection or to identify H. pylori
CC   polypeptide binding compounds, useful as potential H. pylori life cycle
CC   activators or inhibitors. The DNA and probes derived from it may be
CC   used for the identification of H. pylori in a sample and the diagnosis
CC   of H. pylori infection. Nucleic acid sequences complementary to the DNA
CC   act as antisense sequences and can be used to prevent the translation of
CC   H. pylori mRNA. Antibodies against the protein can be used in
CC   immunoassays to evaluate the abundance and distribution of
CC   H. pylori-specific antigens. The genomic sequence of H. pylori
CC   (ATCC 55679) was determined from overlapping contigs generated by
CC   mechanically shearing the bacterial DNA. The sequences were analysed
CC   for ORF of at least 180 nucleotides, and the predicted coding regions
CC   defined by computer evaluation. To identify likely H. pylori antigens for
CC   vaccine development, the amino acid sequences predicted from various ORF
CC   were analysed for significant homology to other known or exported
CC   membrane proteins. Having identified and determined the sequences of
CC   interest, particular regions can be isolated from H. pylori by PCR
CC   amplification for recombinant polypeptide production, e.g. in E. coli
CC   hosts.
SQ   Sequence   307 AA;

 W55590  Length: 307  February 11, 2000 15:48  Type: P  Check: 428  ..

       1  MKKVLLLTLS LSLSFWLHAE RNGFYLGLNF AEGSYIQGQG SIGEKASAEN

      51  ALNQAINNAK NSLFPNTKAI RDVQNALNAV KDSNKIANRF AGNGGSGGIF

     101  NELSLGYKYF LGKKRIIGFR HSLFFGYQLG GVGSVPGSGL IAFLPYGFNT

     151  DLLINWTNDK RASQEYVERR VKGLSIFYKD MTGRTLDADT LKRASRHIIR

     201  KSSGLVIGME LGASTWFASN NLTPFNQVKS RTIFQLQGKF GVRFSSDEYD

     251  IDRYGDENYL GGSSVELGVK VPAFKVNYYS DDYGDKLDYK RVVSVYLNYT

     301  YNFKNKH

!!AA_SEQUENCE 1.0
ID   W55815 standard; Protein; 271 AA.
```

AC W55815;
DT 15-JUL-1998 (first entry)
DE Streptomyces roseofulvus frenolicin gene cluster protein 16.
KW Streptomyces roseofulvus; frenolicin gene cluster; frenolicin B;
KW antibiotic.
OS Streptomyces roseofulvus.
PN J10094395-A.
PD 14-APR-1998.
PF 07-MAY-1997; 116652.
PR 04-APR-1997; US-042935.
PR 07-MAY-1996; US-016753.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
DR WPI; 98-279231/25.
DR N-PSDB; V25925.
PT Frenolicin gene cluster - useful as an antibiotic
PS Claim 8; Page 30; 50pp; Japanese.
CC The present sequence represents a protein from the frenolicin gene
CC cluster from Streptomyces roseofulvus. The present invention describes:
CC (1) a vector containing the frenolicin gene cluster DNA sequence
CC operably connected to an expression control sequence; (2) a host cell
CC transformed by the above vector; (3) a protein coded by the above DNA
CC sequence; (4) a method for the preparation of frenolicin or a
CC biosynthetic intermediate for it in which the above cell is cultured
CC and frenolicin or its biosynthetic intermediate is isolated from the
CC culture or the cell; (5) a method for the preparation of frenolicin B
CC by oxidising frenolicin, and (6) a method for the preparation of a
CC feed composition by mixing frenolicin with other components.
CC Frenolicin B is useful as an antibiotic.
SQ Sequence 271 AA;

W55815 Length: 271 February 11, 2000 15:48 Type: P Check: 5054 ..

1 MTTAAPHTRP GEAGTTRGPA LVTGATRGIG LAVAEALVAR GYPVVVCARD

51 AEAVARTVKE LAAGGARVEG VVADVTDAAS VHELVATTVA RFGPVEVLVN

101 NAGRSGGGVT AELSESLWDD VIATNLKSVF LVTREVLTTG GMTGRGRGVV

151 NIASTGGKQG VVFGAPYSAS KHGVVGFTKA LGLELARSGI TVNAVCPGYV

201 ETPMAAGVRR HYADLWDVTE EDVLARFEAK IPLGRYTRPD EVAALVDYLV

251 TDAAAAVTAQ ALNVCGGLGN Y

!!AA_SEQUENCE 1.0
ID W56137 standard; Peptide; 186 AA.
AC W56137;
DT 14-JUL-1998 (first entry)
DE Open reading frame 2 peptide of bacteriophage phi-CPG1.
KW Circular; bacteriophage phi-CPG1; avian C. psittaci; Chp1;
KW Chlamydia psittaci strain Guinea pig Inclusion Conjunctivitis; GPIC;
KW blindness; ocular chronic disease trachoma; vaccine.
OS Bacteriophage phi-CPG1.
PN US5741697-A.
PD 21-APR-1998.
PF 30-NOV-1995; 565386.
PR 30-NOV-1995; US-565386.
PA (UYRP) UNIV ROCHESTER.
PI Bavoil PM, Hsia R;
DR WPI; 98-260507/23.
DR N-PSDB; V16865.
PT Bacteriophage phi-CPG1 genomic DNA - is useful for transforming
PT mammalian Chlamydia strains
PS Claim 12; Columns 23-24; 28pp; English.
CC The present sequence represents the product of the second open reading
CC frame of a circular bacteriophage, designated phi-CPG1. This
CC bacteriophage is able transfect Chlamydia psittaci strain Guinea pig
CC Inclusion Conjunctivitis (GPIC). Phi-CPG1 was identified by homology to
CC a previously characterised Chlamydia bacteriophage of avian C. psittaci
CC named Chp1. Chlamydia strains can be responsible for blindness through
CC ocular chronic disease trachoma. Phi-CPG1 can be used as a cloning
CC vector, especially as a means to transfer mutant or foreign genes into
CC Chlamydia. Phi-CPG1 can also be used to develop Chlamydia vaccines,
CC diagnostic tests and the study for Chlamydia infection and therapeutic
CC reagents and strategies in infections.
SQ Sequence 186 AA;

W56137 Length: 186 February 11, 2000 15:48 Type: P Check: 7408 ..

1 MNPEQLTNTL GSAVSGVAQG LSFLPGIASG VLGYLGAQKQ NATAKQIARE

51 QMAFQERMSN TAYQRAMEDM KKAGLNPMLA FSKGGASSPA GASWSPNNPV

101 ESAMNSGLAV QRLTYERKKM QAELQNLREQ NRLIRNQAIR EGYLAERDKY

151 MRVAGVPVAT EMLDRTSGLL SSSAKAFKNL FSRKGR

!!AA_SEQUENCE 1.0
ID W51011 standard; protein; 277 AA.
AC W51011;
DT 10-AUG-1998 (first entry)
DE Human liver carbonyl reductase.
KW Human carbonyl reductase; HCRD; inflammatory disorder; cancer;
KW immunological disorder; prostaglandin E.
OS Homo sapiens.
PN US5756299-A.
PD 26-MAY-1998.
PF 09-DEC-1996; 762129.
PR 09-DEC-1996; US-762129.
PA (INCY-) INCYTE PHARM INC.
PI Goli SK, Hillman JL;
DR WPI; 98-321527/28.
PT DNA encoding human carbonyl reductase polypeptide - useful for
PT recombinant production of the enzyme and anti:sense treatment of
PT diseases associated with the enzyme
PS Disclosure; Figure 2; 27pp; English.
CC The invention relates to DNA encoding a human carbonyl reductase protein
CC (HCRD) sequence. Also claimed are: a probe that hybridises to the DNA
CC sequence; an expression vector containing the DNA sequence; and a host
CC cell containing the vector. The enzyme catalyses the reduction of
CC carbonyls to alcohols, physiologically functioning as dehydrogenases of
CC prostaglandin or hydroxysteroids, as well as in drug metabolism. The cell
CC can be used to produce a recombinant human carbonyl reductase protein,
CC which can be used to treat inflammatory or immunological disorders
CC resulting from excessive prostaglandin E production, or to prevent cancer
CC development after exposure to carcinogens. The vector can be used for the
CC same purpose. The probe can be used to detect carbonyl reductase nucleic
CC acids in biological samples. Antisense constructs based on the
CC polynucleotides can be used to control expression of the gene. The
CC present sequence represents human liver carbonyl reductase.
SQ Sequence 277 AA;

W51011 Length: 277 February 11, 2000 15:48 Type: P Check: 7216 ..

1 MSSGIHVALV TGGNKGIGLA IVRDLCRLFS GDVVLTARDV TRGQAAVQQL

51 QAEGLSPRFH QLDIDDLQSI RALRDFLRKE YGGLDVLVNN AGIAFKVADP

101 TPFHIQAEVT MKTNFFGTRD VCTELLPLIK PQGRVVNVSS IMSVRALKSC

151 SPELQQKFRS ETITEEELVG LMNKFVEDTK KGVHQKEGWP SSAYGVTKIG

201 VTVLSRIHAR KLSEQRKGDK ILLNACCPGW VRTDMAGPKA TKSPEEGAET

251 PVYLALLPPD AEGPHGQFVS EKRVEQW

!!AA_SEQUENCE 1.0
ID W28052 standard; Protein; 269 AA.
AC W28052;
DT 27-AUG-1998 (first entry)
DE Amino acid sequence of phospho-beta-glucosidase (bglA).
KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome.
OS Staphylococcus aureus.

```
FH   Key             Location/Qualifiers
FT   Misc_difference 1. .269
FT                   /note= "residues designated X are not defined in
FT                        the specification"
PN   WO9730070-A1.
PD   21-AUG-1997.
PF   19-FEB-1997; U02318.
PR   20-FEB-1996; US-011888.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PI   Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,
PI   Pratt JM, Reichard RW, Rosenberg M, Ward JM;
DR   WPI; 97-424969/39.
DR   N-PSDB; T84002.
PT   Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
PT   to isolate antimicrobial compounds, and in vaccines against S.
PT   aureus infection
PS   Claim 6; Pages 433-434; 989pp; English.
CC   The present sequence represents a Staphylococcus aureus protein, that,
CC   based on homology with a Bacillus subtilis protein, is believed to be
CC   a phospho-beta-glucosidase (bglA). The DNA sequence was isolated from
CC   a library of clones of S. aureus WCUH 29 in Escherichia coli. The DNA
CC   sequence can be used in the construction of ribozymes and antisense
CC   sequences to control the expression of Staphylococcal genes. The DNA
CC   sequence is also useful as a source of regulatory elements for the
CC   control of bacterial gene expression. The present protein may be used
CC   to produce vaccines to enable a host to produce specific antibodies
CC   with antibacterial action. These vaccines and antibodies would protect
CC   a host against invasion by S. aureus, and conditions relating to
CC   Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC   skin syndrome, and toxic shock syndrome.
SQ   Sequence   269 AA;

 W28052  Length: 269  February 11, 2000 15:48  Type: P  Check: 2273  ..

       1  MGLKCLRTSI AWTRIFPNGD EDVPNEEGLA FYDRIFDELI AQGIEPVVTL

      51  SHFEMPLHLA KHYGGFRNRE VVDYFVHFAR VVFERYKDKV TYWMTFNEIN

     101  NQMDTSNPIF LWTNSGVALT ENDNPEEVLY QVAHHELLAS ALAVRLGKEI

     151  NPKFKIGTMI SHVPIYPYSC HPKDMMEAQI ANRLRFFFPD VQVRGYYPSY

     201  AKKMLARKGY DVGWQEGDDS ILQQGTXDYI GFSYYMSTAV KHDVDTTXEN

     251  NIVNGGFGFI CGESAYRNE

!!AA_SEQUENCE 1.0
ID   W37816 standard; Protein; 317 AA.
AC   W37816;
DT   28-AUG-1998  (first entry)
DE   Human secreted apoptosis-related protein hSARP3.
KW   Secreted apoptosis-related protein; SARP; hSARP3; human;
KW   prostate cancer; breast cancer; diagnosis; gene therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1. .28
FT                   /label= Sig_peptide
FT   Protein         29. .317
FT                   /label= Mat_protein
PN   WO9813493-A2.
PD   02-APR-1998.
PF   24-SEP-1997; U17154.
PR   11-OCT-1996; US-028363.
PR   24-SEP-1996; US-026603.
PA   (LXRB-) LXR BIOTECHNOLOGY INC.
PI   Melkonyan H, Umansky S;
DR   WPI; 98-230704/20.
DR   N-PSDB; V19114.
PT   New secreted apoptosis-related proteins - useful for modulating
PT   apoptosis, particularly for treatment of prostatic or breast cancer,
PT   also for diagnosis and monitoring of disease
PS   Claim 1; Page 56-58; 101pp; English.
CC   This polypeptide comprises human secreted apoptosis-related
CC   protein hSARP3 that modulates apoptosis through cell-cell or
CC   cell-extracellular matrix signalling.  Its amino acid sequence was
CC   deduced from a clone (see V19114) obtained from a human pancreas
CC   cDNA library.  Murine mSARP1 (see W37814), human hSARP1 (see
CC   W37816) and hSARP2 (see W37817) proteins are also claimed.  hSARP3
CC   is expressed predominantly in pancreas.  SARP polypeptides can be
CC   obtained from recombinant host cells.  Antibodies specific for SARP
CC   polypeptides can be used in immunoassays for detecting levels of
CC   expression of SARP, particularly for diagnosis or monitoring of
CC   diseases associated with SARP expression.  Specifically, they are
CC   used to detect cancer of the prostate or breast (by detecting
CC   hSARP1 and 2, respectively).  SARP polypeptides and nucleic acids
CC   can also be used to treat these cancers, or more generally
CC   apoptosis-related disease (e.g. infection with HIV or reperfusion
CC   injury), also (not claimed) to prevent apoptosis in cultured cells,
CC   to improve preservation of organs for transplantation, for in situ
CC   preservation for by-pass operations and to treat dermatological
CC   disorders.  SARP polypeptides can also be used to identify agents,
CC   potentially useful therapeutically, that modulate the effects of
CC   SARP on Wnt-frizzled protein interaction.
SQ   Sequence   317 AA;

 W37816  Length: 317  February 11, 2000 15:48  Type: P  Check: 3813  ..

       1  MRAAAAAGGV RTAALALLLG ALHWAPARCE EYDYYGWQAE PLHGRSYSKP

      51  PQCLDIPADL PLCHTVGYKR MRLPNLLEHE SLAEVKQQAS SWLPLLAKRC

     101  HSDTQVFLCS LFAPVCLDRP IYPCRSLCEA VRAGCAPLME AYGFPWPEML

     151  HCHKFPLDND LCIAVQFGHL PATAPPVTKI CAQCEMEHSA DGLMEQMCSS

     201  DFVVKMRIKE IKIENGDRKL IGAQKKKKLL KPGPLKRKDT KRLVLHMKNG

     251  AGCPCPQLDS LAGSFLVMGR KVDGQLLLMA VYRWDKKNKE MKFAVKFMFS

     301  YPCSLYYPFF YGAAEPH

!!AA_SEQUENCE 1.0
ID   W59024 standard; Protein; 71 AA.
AC   W59024;
DT   14-SEP-1998  (first entry)
DE   Enterocin-900.
KW   Enterocin-900; bacteriocin; signal peptide; protein secretion;
KW   preservative; food spoilage; lactic acid bacterium; antibacterial.
OS   Enterococcus faecium 900.
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= Sig_peptide
FT   Protein         19. .71
FT                   /label= Mat_protein
PN   WO9809639-A1.
PD   12-MAR-1998.
PF   05-SEP-1997; U15609.
PR   05-SEP-1996; US-026257.
PA   (UYAL-) UNIV ALBERTA.
PI   Franz CM, Greer GG, Leisner JJ, McCormick JK, McMullen LM,
PI   Pkns J, Poona, Stiles ME, Van Belkum MJ, Vederas JC,
PI   Worobo RJ, Worobo RW;
DR   WPI; 98-193319/17.
DR   N-PSDB; V11704-05.
PT   Bacterial growth inhibiting peptide(s) enterocin 900 or brochocin C
PT   - used to inhibit growth of susceptible bacteria in e.g. foodstuff,
PT   living animal, food preparation area or fermentation vessel
PS   Example 7; Page 138-139; 174pp; English.
CC   This polypeptide comprises the enterocin-900 pre-protein of
CC   Enterococcus faecium 900.  Enterocin-900 is a novel broad
CC   spectrum bacteriocin that has activity against other strains of
CC   Enterococcus as well as many other organisms.  The invention
CC   includes a method for inhibiting pathogenic bacteria by providing
CC   a bacteriocin selected from enterocin-900 and brochocin-C (see
CC   W59021), either as a composition or by providing a bacterial
CC   source of the bacteriocins.  This is used to inhibit spoilage
```

CC bacteria in foodstuffs, such as meat, inhibit pathogenic bacteria
CC topically in animals, e.g. to treat mastitis (claimed), and inhibit
CC bacterial infection of fermentation reactors.
SQ Sequence 71 AA;

W59024 Length: 71 February 11, 2000 15:48 Type: P Check: 9911 ..

1 MQNVKELSTK EMKQIIGGEN DHRMPNELNT PNNLSKGGAK CGAAIAGGLF

51 GIPKGPLAWA AALANVYSKC N

!!AA_SEQUENCE 1.0
ID W57330 standard; Protein; 339 AA.
AC W57330;
DT 14-SEP-1998 (first entry)
DE Glycerol-3-phosphate dehydrogenase gpsA.
KW Glycerol-3-phosphate dehydrogenase; G3PDH; gpsA; yeast.
OS Saccharomyces sp.
PN WO9821340-A1.
PD 22-MAY-1998.
PF 10-NOV-1997; U20293.
PR 13-NOV-1996; US-030602.
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PA (GEMV) GENENCOR INT INC.
PI Bulthuis BA, Gatenby AA, Haynie SL, Hsu AK, Lareau RD;
DR WPI; 98-297943/26.
PT Fermentative production of glycerol using recombinant host -
PT containing genes for glycerol-3-phosphate dehydrogenase and-or
PT glycerol-3- phosphatase
PS Claim 9; Page 36-37; 57pp; English.
CC This claimed Saccharomyces polypeptide comprises a
CC glycerol-3-phosphate dehydrogenase (G3PDH) that catalyses the
CC conversion of dihydroxyacetone phosphate to glycerol-3-phosphate.
CC It is encoded by the gpsA gene. The invention provides recombinant
CC organisms that express G3PDH and/or glycerol-3-phosphatase (G3P)
CC (see also W57324-32) useful for the production of glycerol from a
CC variety of C-sources. A host cell is preferably transformed with a
CC cassette containing either a G3PDH gene and/or a G3P gene and then
CC cultured in the presence of a mono-, oligo-, polysaccharide or
CC 1C-substrate. The glycerol obtained is used in cosmetics, liquid
CC soaps, pharmaceuticals, lubricants and antifreezes; its esters are
CC used in the oil and fat industries. The method produces glycerol
CC rapidly and inexpensively without generation of polluting
CC by-products.
SQ Sequence 339 AA;

W57330 Length: 339 February 11, 2000 15:48 Type: P Check: 4443 ..

1 MNQRNASMTV IGAGSYGTAL AITLARNGHE VVLWGHDPEH IATLERDRCN

51 AAFLPDVPFP DTLHLESDLA TALAASRNIL VVVPSHVFGE VLRQIKPLMR

101 PDARLVWATK GLEAETGRLL QDVAREALGD QIPLAVISGP TFAKELAAGL

151 PTAISLASTD QTFADDLQQL LHCGKSFRVY SNPDFIGVQL GGAVKNVIAI

201 GAGMSDGIGF GANARTALIT RGLAEMSRLG AALGADPATF MGMAGLGDLV

251 LTCTDNQSRN RRFGMMLGQG MDVQSAQEKI GQVVEGYRNT KEVRELAHRF

301 GVEMPITEEI YQVLYCGKNA REAALTLLGR ARKDERSSH

!!AA_SEQUENCE 1.0
ID W62641 standard; Protein; 137 AA.
AC W62641;
DT 18-SEP-1998 (first entry)
DE Flea serine protease inhibitor variable domain protein pfSPI14-137.
KW Serine protease inhibitor; SPI; anti-haematophagous ectoparasite;
KW vaccine; control; flea; acarina; immune response.
OS Ctenocephalides felis.
PN WO9820034-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20678.
PR 07-NOV-1996; US-745995.
PA (HESK-) HESKA CORP.
PI Brandt KS, Maddux JD, Silver GM, Wisnewski N;
DR WPI; 98-286864/25.
DR N-PSDB; V38773.
PT New nucleic acid encoding new serine protease inhibitors from fleas
PT - and related antibodies, inhibitors, recombinant viruses and cells,
PT useful for control of fleas on animals, especially cats and dogs
PS Claim 4; Page 151; 176pp; English.
CC The present sequence, designated pfSPI14-137, represents Ctenocephalides
CC felis serine protease inhibitor (SPI) variable domain protein. The SPI
CC protein represents a novel target for anti-haematophagous ectoparasitic
CC vaccines and drugs. Compositions containing the inhibitor are used to
CC control fleas (adults or larvae) on mammals or birds, particularly cats
CC and dogs. More generally the composition is also used for control of
CC insects and acarina, on animals or on stored goods, plants, trees etc.
CC including domestic pests and insects that are vectors of disease. The SPI
CC proteins are able to induce a protective immune response against SPI.
SQ Sequence 137 AA;

W62641 Length: 137 February 11, 2000 15:48 Type: P Check: 7060 ..

1 EEKLQNVDLQ NLTQRMYSVE VILDLPKFKI ESEINLNDPL KKLGMSDMFV

51 PGKADFKGLL EGSDEMLYIS KVIQKAFIEV NEEGAEAAAA TAVFATRRVI

101 KVLAKEIFNC DHPFYFALVH SQEGTSAPLF TGAFRTP

!!AA_SEQUENCE 1.0
ID W55052 standard; Protein; 118 AA.
AC W55052;
DT 28-SEP-1998 (first entry)
DE Sunflower antifungal protein MS59 fragment.
KW Antifungal; fungicide; MS59; sunflower; carbohydrate oxidase;
KW glucose oxidase; transgenic plant; Phytophthora; Pythium;
KW crop protection; disease resistance.
OS Helianthus annuus cv. zebulon.
FH Key Location/Qualifiers
FT Misc_difference 68
FT /label= Arg, Lys
PN WO9813478-A2.
PD 02-APR-1998.
PF 04-SEP-1997; E04923.
PR 19-MAR-1997; EP-200831.
PR 04-SEP-1996; EP-202466.
PA (MOGE-) MOGEN INT NV.
PI Custers JHHV, Lageweg W, Melchers LS, Ponstein AS,
PI Sela-Buurlage MB, Stuiver MH, Van Deventer-Troost JPE;
DR WPI; 98-230692/20.
DR N-PSDB; V27251.
PT New plant proteins having anti-fungal activity - useful as, e.g.
PT carbohydrate oxidase(s) for protection against Phytophthora and
PT Pythium sp.
PS Claim 5; Page 53-54; 139pp; English.
CC This polypeptide is the deduced product of a DNA sequence (see
CC V27251) amplified from sunflower cv. zebulon genomic DNA using
CC PCR primers (see V27249-50) based on tryptic peptides (see
CC W55050-51) of the isolated sunflower MS59 antifungal protein.
CC A Southern blot of sunflower genomic DNA using this DNA fragment
CC revealed the existence of multiple homologous sequences in the
CC sunflower genome. A full-length clone (see V27260) for sunflower
CC MS59 (see W55053) was subsequently obtained. Claimed antifungal
CC proteins, which may include this MS59 polypeptide sequence, have
CC a mol.wt. of 55-65 kDa (SDS-PAGE), have carbohydrate oxidase
CC (especially glucose oxidase) activity, show anti-Phytophthora
CC and/or anti-Pythium activity, can be expressed in transgenic
CC plants to reduce susceptibility to infection by fungi, or
CC expressed in host cells for use in antifungal compositions. Plants
CC engineered to express the antifungal proteins require reduced
CC treatments with fungicides and have a longer shelf life.
SQ Sequence 118 AA;

W55052 Length: 118 February 11, 2000 15:48 Type: P Check: 4306 ..

```
      1  DPSFPITGEV YTPGNSSFPT VLQNYIRNLR FNETTTPKPF LIITAEHVSH

     51  IQAAVVCGKQ NRLLLKTXSG GHDYEGLSYL TNTNQPFFIV DMFNLRSINV

    101  DIEQETAWVQ AGATLGEV

!!AA_SEQUENCE 1.0
ID   W48722 standard; Protein; 355 AA.
AC   W48722;
DT   25-SEP-1998  (first entry)
DE   Human V28 seven transmembrane receptor.
KW   V28; placenta; seven transmembrane receptor; 7TM; signal transduction;
KW   immunology; inflammation.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Domain          26. .56
FT                   /note= "Transmembrane domain 1"
FT   Domain          68. .92
FT                   /note= "Transmembrane domain 2"
FT   Domain          107. .125
FT                   /note= "Transmembrane domain 3"
FT   Domain          146. .167
FT                   /note= "Transmembrane domain 4"
FT   Domain          197. .219
FT                   /note= "Transmembrane domain 5"
FT   Domain          232. .253
FT                   /note= "Transmembrane domain 6"
FT   Domain          273. .297
FT                   /note= "Transmembrane domain 7"
PN   US5759804-A.
PD   02-JUN-1998.
PF   17-NOV-1993; 153848.
PR   17-NOV-1992; US-977452.
PA   (ICOS-) ICOS CORP.
PI   Godiska R, Gray PW, Schweickart VL;
DR   WPI; 98-332132/29.
DR   N-PSDB; V18343, V18352.
PT   DNA encoding V28 seven transmembrane receptor polypeptide - useful
PT   for producing recombinant polypeptide and anti-V28 antibodies, and
PT   in screening assays for V28 agonists and antagonists
PS   Claim 1; Columns 61-64; 56pp; English.
CC   The present sequence claimed by the invention represents the human
CC   V28 seven transmembrane (7TM) receptor.  It is encoded by the full
CC   length V28 genomic DNA (V18343) or by the V28 cDNA (V18352).  As
CC   the V28 protein (W48722) is a 7TM receptor it is probably involved in
CC   signal transduction.  The invention also claims that cells transformed
CC   with V28 DNA can be used to produce the recombinant polypeptide, to
CC   produce anti-V28 antibodies or in screening assays for V28 agonists or
CC   antagonists.  The antibodies, agonists and antagonists could then be
CC   used to modulate V28 receptor-ligand binding, for e.g. in immunological
CC   and/or inflammatory events in vivo.
SQ   Sequence   355 AA;

 W48722  Length: 355  February 11, 2000 15:48  Type: P  Check: 2863  ..

      1  MDQFPESVTE NFEYDDLAEA CYIGDIVVFG TVFLSIFYSV IFAIGLVGNL

     51  LVVFALTNSK KPKSVTDIYL LNLALSDLLF VATLPFWTHY LINEKGLHNA

    101  MCKFTTAFFF IGFFGSIFFI TVISIDRYLA IVLAANSMNN RTVQHGVTIS

    151  LGVWAAAILV AAPQFMFTKQ KENECLGDYP EVLQEIWPVL RNVETNFLGF

    201  LLPLLIMSYC YFRIIQTLFS CKNHKKAKAI KLILLVVIVF FLFWTPYNVM

    251  IFLETLKLYD FFPSCDMRKD LRLALSVTET VAFSHCCLNP LIYAFAGEKF

    301  RRYLYHLYGK CLAVLCGRSV HVDFSSSESQ RSRHGSVLSS NFTYHTSDGD

    351  ALLLL

!!AA_SEQUENCE 1.0
ID   W54728 standard; peptide; 24 AA.
AC   W54728;
DT   25-SEP-1998  (first entry)
DE   Peptide from Ii p80-103.
KW   Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;
KW   vaccine; treatment.
OS   Synthetic.
PN   WO9813378-A1.
PD   02-APR-1998.
PF   25-SEP-1997; NL0536.
PR   26-SEP-1996; EP-202701.
PA   (UYLE-) RIJKSUNIV LEIDEN.
PI   Drijfhout JW, Koning F;
DR   WPI; 98-230631/20.
PT   Increasing uptake and presentation of antigen(s) - by adding mannose
PT   residue(s) to antigen for increasing T cell response, useful in,
PT   e.g. vaccines against viral infection(s)
PS   Disclosure; Page 34; 47pp; English.
CC   The peptides W54559-W54809 are examples of peptides to which at least 1
CC   (preferably 2) mannose can be attached to increase their uptake as
CC   antigens by antigen-presenting cells.  Uptake of agonist mannosylated
CC   peptides will increase the T cell response, whereas uptake of antagonist
CC   peptides blocks the T cell response.  Blocking binding of immunogenic
CC   autoantigens can be used in treatment of type I diabetes, rheumatoid
CC   arthritis, graft rejection etc., also to induce T-cell non-
CC   responsiveness.  Vaccines containing mannosylated antigen are used to
CC   prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths
CC   and parasites.
SQ   Sequence   24 AA;

 W54728  Length: 24  February 11, 2000 15:48  Type: P  Check: 3126  ..

      1  LPKPPKPVSK MRMATPLLMQ ALPM

!!AA_SEQUENCE 1.0
ID   W60258 standard; Protein; 339 AA.
AC   W60258;
DT   28-SEP-1998  (first entry)
DE   Klebsiella pneumoniae glycerol-3-phosphate dehydrogenase.
KW   glycerol-3-phosphate dehydrogenase; production;
KW   1,3-propanediol; recombinant.
OS   Klebsiella pneumoniae.
PN   WO9821341-A2.
PD   22-MAY-1998.
PF   13-NOV-1997; U20873.
PR   13-NOV-1996; US-030601.
PA   (GEMV ) GENENCOR INT INC.
PI   Chase MW, Diaz-torres M, Dunn-coleman NS, Trimbur D;
DR   WPI; 98-297944/26.
PT   New method for increasing production of 1,3-propane:diol - comprises
PT   fermentation of inexpensive carbon sources by microorganism
PT   expressing dehydratase, used, e.g. to prolong half-life of enzyme
PS   Disclosure; Page 70-71; 133pp; English.
CC   The sequence is that of glycerol-3-phosphate dehydrogenase.
CC   It was used as part of a method of fermentative production
CC   of 1,3-propanediol (1,3-pd), using an organism comprising
CC   at least 1 gene encoding a dehydratase, is improved by
CC   inserting into the host a gene encoding protein X and culturing
CC   the transformant in presence of a carbon source (e.g. mono-, oligo-
CC   or poly-saccharide or 1C substrate) convertible to 1,3-pd.
CC   1,3-pd is a starting material for polyesters, polyurethanes and
CC   cyclic compounds. 1,3-pd can now be produced by a single
CC   recombinant organism from inexpensive carbon sources such
CC   as glucose (rather than costly glycerol or dihydroxyacetone),
CC   rapidly and without causing pollution.
SQ   Sequence   339 AA;

 W60258  Length: 339  February 11, 2000 15:48  Type: P  Check: 4443  ..

      1  MNQRNASMTV IGAGSYGTAL AITLARNGHE VVLWGHDPEH IATLERDRCN

     51  AAFLPDVPFP DTLHLESDLA TALAASRNIL VVVPSHVFGE VLRQIKPLMR
```

```
     101  PDARLVWATK GLEAETGRLL QDVAREALGD QIPLAVISGP TFAKELAAGL

     151  PTAISLASTD QTFADDLQQL LHCGKSFRVY SNPDFIGVQL GGAVKNVIAI

     201  GAGMSDGIGF GANARTALIT RGLAEMSRLG AALGADPATF MGMAGLGDLV

     251  LTCTDNQSRN RRFGMMLGQG MDVQSAQEKI GQVVEGYRNT KEVRELAHRF

     301  GVEMPITEEI YQVLYCGKNA REAALTLLGR ARKDERSSH

!!AA_SEQUENCE 1.0
ID   W60267 standard; Protein; 387 AA.
AC   W60267;
DT   28-SEP-1998  (first entry)
DE   Klebsiella pneumoniae DHAT protein.
KW   glycerol-3-phosphate dehydrogenase; production;
KW   1,3-propanediol; recombinant.
OS   Klebsiella pneumoniae.
PN   WO9821341-A2.
PD   22-MAY-1998.
PF   13-NOV-1997; U20873.
PR   13-NOV-1996; US-030601.
PA   (GEMV ) GENENCOR INT INC.
PI   Chase MW, Diaz-torres M, Dunn-coleman NS, Trimbur D;
DR   WPI; 98-297944/26.
PT   New method for increasing production of 1,3-propane:diol - comprises
PT   fermentation of inexpensive carbon sources by microorganism
PT   expressing dehydratase, used, e.g. to prolong half-life of enzyme
PS   Disclosure; Page 96-97; 133pp; English.
CC   The sequence is that of DHAT protein used as part of a method of
CC   fermentative production of 1,3-propanediol (1,3-pd), using an organism
CC   comprising at least 1 gene encoding a dehydratase, is improved by
CC   inserting into the host a gene encoding protein X and culturing
CC   the transformant in presence of a carbon source (e.g. mono-, oligo-
CC   or poly-saccharide or 1C substrate) convertible to 1,3-pd.
CC   1,3-pd is a starting material for polyesters, polyurethanes and
CC   cyclic compounds. 1,3-pd can now be produced by a single
CC   recombinant organism from inexpensive carbon sources such
CC   as glucose (rather than costly glycerol or dihydroxyacetone),
CC   rapidly and without causing pollution.
SQ   Sequence   387 AA;

 W60267  Length: 387  February 11, 2000 15:48  Type: P  Check: 843  ..

       1  MSYRMFDYLV PNVNFFGPNA ISVVGERCQL LGGKKALLVT DKGLRAIKDG

      51  AVDKTLHYLR EAGIEVAIFD GVEPNPKDTN VRDGLAVFRR EQCDIIVTVG

     101  GGSPHDCGKG IGIAATHEGD LYQYAGIETL TNPLPPIVAV NTTAGTASEV

     151  TRHCVLTNTE TKVKFVIVSW RKLPSVSIND PLLMIGKPAA LTAATGMDAL

     201  THAVEAYISK DANPVTDAAA MQAIRLIARN LRQAVALGSN LQARENMAYA

     251  SLLAGMAFNN ANLGYVHAMA HQLGGLYDMP HGVANAVLLP HVARYNLIAN

     301  PEKFADIAEL MGENITGLST LDAAEKAIAA ITRLSMDIGI PQHLRDLGVK

     351  EADFPYMAEM ALKDGNAFSN PRKGNEQEIA AIFRQAF

!!AA_SEQUENCE 1.0
ID   W60764 standard; Protein; 260 AA.
AC   W60764;
DT   28-SEP-1998  (first entry)
DE   Rainbow trout interleukin 1 beta.
KW   interleukin 1 beta; IL1; beta; vaccine; fish; medicine.
OS   Oncorhynchus mykiss.
PN   WO9817802-A1.
PD   30-APR-1998.
PF   16-OCT-1997; G02855.
PR   06-NOV-1996; GB-023173.
PR   17-OCT-1996; GB-021681.
PA   (UYAB-) UNIV ABERDEEN.
PI   Cunningham C, Secombes CJ, Zou J;
DR   WPI; 98-261498/23.
DR   N-PSDB; V36179.
PT   Rainbow Trout Interleukin-1beta - can be administered as adjuvant in
PT   existing and developing fish vaccines
PS   Claim 14; Fig 3; 46pp; English.
CC   The sequence is that of rainbow trout interleukin 1 (IL-1)
CC   beta which can be used in the manufacture of a medicament for fish
CC   e.g. it can be used as a adjuvant in existing and developing vaccines.
SQ   Sequence   260 AA;

 W60764  Length: 260  February 11, 2000 15:48  Type: P  Check: 7422  ..

       1  MDFESNYSLI KNTSESAAWS SKLPQGLDLE VSHHPITMRH IANLIIAMER

      51  LKGGEGVTMG TEFKDKDLLN FLLESAVEEH IVLELESAPP ASRRAAGFSS

     101  TSQYECSVTD SENKCWVLMN EAMELHAMML QGGSSYHKVH LNLSSYVTPV

     151  PIETEARPVA LGIKGSNLYL SCSKSGGRPT LHLEEVADKD QLKSISQQSD

     201  MVRFLFYRRN TGVDISTLES ASFRNWFIST DMQQDYTKPV DMCQKAAPNR

     251  LTTFTIQRHN

!!AA_SEQUENCE 1.0
ID   W57889 standard; Protein; 271 AA.
AC   W57889;
DT   23-SEP-1998  (first entry)
DE   Corn raffinose synthetase.
KW   Raffinose synthetase; metabolism modification; food additive;
KW   gastrointestinal flora; corn.
OS   Zea mays.
PN   EP-849359-A2.
PD   24-JUN-1998.
PF   18-DEC-1997; 122417.
PR   18-DEC-1996; JP-338673.
PA   (SUMO ) SUMITOMO CHEM CO LTD.
PI   Oeda K, Wantanabe E;
DR   WPI; 98-324670/29.
DR   N-PSDB; V40803.
PT   New nucleic acid molecule encoding plant raffinose synthetase -
PT   capable of producing raffinose, used as food additives with
PT   beneficial effects on gastrointestinal flora
PS   Claim 1; Page 39-40; 44pp; English.
CC   This sequence represents the corn raffinose synthetase of the
CC   invention. The raffinose synthetase is capable of producing raffinose by
CC   combining a D-galactosyl group through an alpha (1-6) bond with a
CC   hydroxyl group attached to the carbon atom at position 6 of a D-glucose
CC   residue in a sucrose molecule. The DNA can be used to modify metabolism
CC   of a host organism by introducing into the host organism or its cell so
CC   that the content of the raffinose family oligosaccharides in the host
CC   organism or cell is changed. Raffinose oligosaccharides are useful as
CC   food additives with beneficial effects on the gastrointestinal flora.
SQ   Sequence   271 AA;

 W57889  Length: 271  February 11, 2000 15:48  Type: P  Check: 5290  ..

       1  QSTHPCAAFH AASRAISGGP IYVSDSVGQH DFALLRRLAL PDGTVLRCEG

      51  HALPTRDCLF ADPLHDGRTV LKIWNVNRFA GVVGAFNCQG GGWSPEARRN

     101  KCFSEFSVPL AARASPSDVE WKSGKAGPGV SVKDVSQFAV YAVEARTLQL

     151  LRPDEGVDLT LQPFTYELFV VAPVRVISHE RAIKFAPIGL ANMLNTAGAV

     201  QAFEAKKDAS GVTAEVFVKG AGELVAYSSA TPRLCKVNGD EAEFTYKDGV

     251  VTVDVPWSGS SSKLCCVQYV Y

!!AA_SEQUENCE 1.0
ID   W44905 standard; Protein; 39 AA.
AC   W44905;
```

```
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide Delta-PRO4-beta.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30424.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 36; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   39 AA;

 W44905  Length: 39  February 11, 2000 15:48  Type: P  Check: 1135  ..

       1  AAADPITMFS LTRQVLNVGP RVPIGPNPVL PDQRLPSSA

!!AA_SEQUENCE 1.0
ID   W44906 standard; Protein; 70 AA.
AC   W44906;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide Delta-PRO4-int.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30425.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 37; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   70 AA;

 W44906  Length: 70  February 11, 2000 15:48  Type: P  Check: 6449  ..

       1  AAANPLVLEF TDAGKKANWD GPKSWGLRLY RTGTDPITMF SLTRQVLNVG

      51  PRVPIGPNPV LPDQRLPSSA

!!AA_SEQUENCE 1.0
ID   W44907 standard; Protein; 127 AA.
AC   W44907;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide Delta-PRO4-vrb.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30426.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 38; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   127 AA;

 W44907  Length: 127  February 11, 2000 15:48  Type: P  Check: 3740  ..

       1  AAAETTGQAY WKPTSSWDLI SLKRGNTPWD TGCSKVACGP CYDLSKVSNS

      51  FQGATRGGRC NPLVLEFTDA GKKANWDGPK SWGLRLYRTG TDPITMFSLT

     101  RQVLNVGPRV PIGPNPVLPD QRLPSSA

!!AA_SEQUENCE 1.0
ID   W44908 standard; Protein; 79 AA.
AC   W44908;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide PRO-beta.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
```

```
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30427.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 39; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   79 AA;

 W44908  Length: 79  February 11, 2000 15:48  Type: P  Check: 1205  ..

       1  AAADPITMFS LTRQVLNVGP RVPIGPNPVL PDQRLPSSPI EIVPAPQPPS

      51  PLNTSYPPST TSTPSTSPTS PSVPQPPPA

!!AA_SEQUENCE 1.0
ID   W44909 standard; Protein; 110 AA.
AC   W44909;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide PRO-int.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30428.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 40; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   110 AA;

 W44909  Length: 110  February 11, 2000 15:48  Type: P  Check: 4699  ..

       1  AAANPLVLEF TDAGKKANWD GPKSWGLRLY RTGTDPITMF SLTRQVLNVG

      51  PRVPIGPNPV LPDQRLPSSP IEIVPAPQPP SPLNTSYPPS TTSTPSTSPT

     101  SPSVPQPPPA

!!AA_SEQUENCE 1.0
ID   W44910 standard; Protein; 167 AA.
AC   W44910;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide PRO-vrb.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30429.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 41; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   167 AA;

 W44910  Length: 167  February 11, 2000 15:48  Type: P  Check: 1990  ..

       1  AAAETTGQAY WKPTSSWDLI SLKRGNTPWD TGCSKVACGP CYDLSKVSNS

      51  FQGATRGGRC NPLVLEFTDA GKKANWDGPK SWGLRLYRTG TDPITMFSLT

     101  RQVLNVGPRV PIGPNPVLPD QRLPSSPIEI VPAPQPPSPL NTSYPPSTTS

     151  TPSTSPTSPS VPQPPPA

!!AA_SEQUENCE 1.0
ID   W44879 standard; Protein; 104 AA.
AC   W44879;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide PRO+(4070A).
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
```

```
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30398.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 10; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   104 AA;

 W44879  Length: 104  February 11, 2000 15:48  Type: P  Check: 9655  ..

       1  EFTDAGKKAN WDGPKSWGLR LYRTGTDPIT MFSLTRQVLN VGPRVPIGPN

      51  PVLPDQRLPS SPIEIVPAPQ PPSPLNTSYP PSTTSTPSTS PTSPSVPQPP

     101  PAAA

!!AA_SEQUENCE 1.0
ID   W44882 standard; Protein; 61 AA.
AC   W44882;
DT   29-SEP-1998  (first entry)
DE   "polyproline beta-turn helix" polypeptide Delta-PRO+.
KW   Chimeric; polyproline; beta-turn helix; facilitation; inhibition;
KW   retrovirus; envelope glycoprotein; gene therapy; cystic fibrosis;
KW   myopathy; lysosomal disorder; cancer; human immunodeficiency virus.
OS   Synthetic.
PN   WO9744474-A2.
PD   27-NOV-1997.
PF   16-MAY-1997; F00870.
PR   20-MAY-1996; FR-006234.
PA   (CNRS ) CENT NAT RECH SCI.
PI   Cosset FL, Russell SJ, Valsesia S;
DR   WPI; 98-018526/02.
DR   N-PSDB; V30401.
PT   Introducing genes into eukaryotic cells using peptide with two
PT   receptor binding regions - especially as part of a viral
PT   glyco-protein, used in gene therapy to ensure precise targeting of
PT   vectors to particular cell types
PS   Disclosure; Fig 13; 117pp; French.
CC   A novel chimeric protein for gene transfer comprises, at the N-terminus,
CC   a protein binding a target molecule on the cell surface, and, at the
CC   C-terminus, a second protein recognising a secondary target molecule on
CC   the cell surface.  The two domains are separated by a peptide comprising
CC   at least 30% Pro residues arranged so as to induce approximately 180 deg.
CC   folds (beta or reverse turns) in the peptide sequence, the folds being
CC   regularly spread to form a polyproline beta-turn helix structure.  This
CC   sequence is an example of a Pro-rich linking peptide.  The pro-rich
CC   peptide acts by facilitating or inhibiting binding between the C-terminal
CC   protein and its receptor, and inhibition occurs as long as the N-terminal
CC   protein does not interact with its target. The whole construct especially
CC   forms part of a (retro)viral envelope glycoprotein.  The recombinant
CC   (retro)viral particles are used for in vitro or ex vivo transfer of
CC   nucleic acid into target cells, especially for gene therapy, e.g. in
CC   cystic fibrosis, myopathy, lysosomal disorders, cancer, viral (especially
CC   human immunodeficiency virus) infections, etc.
SQ   Sequence   61 AA;

 W44882  Length: 61  February 11, 2000 15:48  Type: P  Check: 8918  ..

       1  EFTDAGKKAN WDGPKSWGLR LYRTGTDPIT MFSLTRQVLN VGPRVPIGPN

      51  PVLPDAAAPH Q

!!AA_SEQUENCE 1.0
ID   W62612 standard; Protein; 369 AA.
AC   W62612;
DT   21-SEP-1998  (first entry)
DE   Human glutamate-binding protein (HGLUBP).
KW   Human; glutamate-binding protein; HGLUBP; Incyte clone 386116;
KW   treatment; prevention; diagnosis; central nervous system disorder;
KW   signalling; glutamate; ischaemia; hypoglycaemia; epilepsy; convulsion;
KW   schizophrenia; neurodegeneration; amyotrophic lateral sclerosis;
KW   lathyrism.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Region          178. .21
FT                   /note= "transmembrane region"
FT   Region          211. .230
FT                   /note= "transmembrane region"
FT   Region          260. .285
FT                   /note= "transmembrane region"
FT   Region          297. .320
FT                   /note= "transmembrane region"
PN   WO9821241-A1.
PD   22-MAY-1998.
PF   10-NOV-1997; U20860.
PR   14-NOV-1996; US-749289.
PA   (INCY-) INCYTE PHARM INC.
PI   Bandman O, Coleman R;
DR   WPI; 98-297868/26.
DR   N-PSDB; V38502.
PT   New human glutamate-binding protein - useful for treatment,
PT   prevention and diagnosis of central nervous system diseases, e.g.
PT   Alzheimer's
PS   Claim 1; Fig 1A-E; 66pp; English.
CC   The present sequence represents human glutamate-binding protein (HGLUBP).
CC   The protein has structural homology to the rat glutamate-binding subunit
CC   of the N-methyl-D-aspartate receptor. The nucleic acid encoding HGLUBP
CC   was first identified in Incyte clone 386116 from a thymus cDNA library.
CC   HGLUBP is useful for treatment, prevention and diagnosis of central
CC   nervous system disorders that require an increase or reduction in
CC   signalling through glutamate, e.g. ischaemia, hypoglycaemia, epilepsy,
CC   convulsions, dementia related to AIDS, schizophrenia, neurodegeneration
CC   such as Alzheimer's, Huntington's, Creutzfeldt-Jakob or Parkinson's
CC   diseases, amyotrophic lateral sclerosis or lathyrism, pain, defective
CC   motor rhythm, vasomotor tone, baroreception and long-term potentiation,
CC   muscle relaxation and sedation. Antibodies can also be used to target
CC   pharmaceuticals to HGLUBP-expressing cells or to diagnose diseases
CC   associated with HGLUBP expression and to monitor treatment.
SQ   Sequence   369 AA;

 W62612  Length: 369  February 11, 2000 15:49  Type: P  Check: 8912  ..

       1  MKRVFWCLGT TILPPTLDIR GGPSHPCPPM LSLPTLGPLT HSPLSSPPPT

      51  VSQGTPMAPA PTPKGATHRV PTPKGATHRG PYPQEGYPQG PYPQGGYPQG

     101  PYPQSPFPPN PYGQPQVFPG QDPDSPQHGN YQEEGPPSYY DNQDFPATNW

     151  DDKSIRQAFI RKVFLVLTLQ LSVTLSTVSV FTFVAEVKGF VRENVWTYYV

     201  SYAVFFISLI VLSCCGDFRR KHPWNLVALS VLTASLSYMV GMIASFYNTE

     251  AVIMAVGITT AVCFTVVIFS MQTRYDFTSC MGVLLVSMVV LFIFAILCIF

     301  IRNRILEIVY ASLGALLFTC FLAVDTQLLL GNKQLSLSPE EYVFAALNLY

     351  TDIINIFLYI LTIIGRAKE
```

```
!!AA_SEQUENCE 1.0
ID   W30679 standard; Protein; 387 AA.
AC   W30679;
DT   12-OCT-1998  (first entry)
DE   1,3-Propanediol oxidoreductase.
KW   1,3-Propanediol oxidoreductase; DHAT; 1,3-propanediol.
OS   Klebsiella pneumoniae ATCC 25955.
PN   WO9821339-A1.
PD   22-MAY-1998.
PF   10-NOV-1997; U20292.
PR   13-NOV-1996; US-030601.
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
PA   (GEMV ) GENENCOR INT INC.
PI   Dias-Torres M, Gatenby AA, Haynie SL, Hsu AK, Lareau RD,
PI   Nagarajan V, Nair RV, Nakamura CE, Payne MS, Picataggio SK,
PI   Trimbur DE, Whited GM;
DR   WPI; 98-297942/26.
DR   N-PSDB; V42015 and V42022.
PT   Fermentative production of 1,3-propanediol  - by single organism
PT   containing cassette comprising specific genes, and capable of using
PT   inexpensive carbon sources
PS   Claim 13; Page 81-82; 95pp; English.
CC   This is the 1,3-propanediol oxidoreductase encoded by the dhaT
CC   gene (see V42015) of Klebsiella pneumoniae ATCC 25955.   This
CC   enzyme catalyses the reduction of 3-hydroxypropionaldehyde to
CC   1,3-propanediol (I).  A claimed method for production of (I)
CC   comprises culturing a microorganism transformed with a cassette
CC   containing at least 1 of the genes (see V42012-21) for
CC   glycerol-3-phosphate dehydrogenase, glycerol-3-phosphatase,
CC   glycerol dehydratase and 1,3-propanediol oxidoreductase (see
CC   W30676-85).  Any of these genes not present on the cassette must
CC   be present endogenously.  The enzyme sequences may include
CC   substitutions, deletions and additions provided activity is not
CC   altered.  A single recombinant organism can now be used for
CC   production of (I) from inexpensive C-sources (contrast use of
CC   glycerol or dihydroxyacetone) without causing pollution.  (I) is
CC   a starting material for polyesters, polyurethanes and cyclic
CC   compounds.
SQ   Sequence   387 AA;

 W30679  Length: 387  February 11, 2000 15:49  Type: P  Check: 843  ..

       1  MSYRMFDYLV PNVNFFGPNA ISVVGERCQL LGGKKALLVT DKGLRAIKDG

      51  AVDKTLHYLR EAGIEVAIFD GVEPNPKDTN VRDGLAVFRR EQCDIIVTVG

     101  GGSPHDCGKG IGIAATHEGD LYQYAGIETL TNPLPPIVAV NTTAGTASEV

     151  TRHCVLTNTE TKVKFVIVSW RKLPSVSIND PLLMIGKPAA LTAATGMDAL

     201  THAVEAYISK DANPVTDAAA MQAIRLIARN LRQAVALGSN LQARENMAYA

     251  SLLAGMAFNN ANLGYVHAMA HQLGGLYDMP HGVANAVLLP HVARYNLIAN

     301  PEKFADIAEL MGENITGLST LDAAEKAIAA ITRLSMDIGI PQHLRDLGVK

     351  EADFPYMAEM ALKDGNAFSN PRKGNEQEIA AIFRQAF

!!AA_SEQUENCE 1.0
ID   W69232 standard; Protein; 263 AA.
AC   W69232;
DT   20-OCT-1998  (first entry)
DE   FcR-II protein sequence.
KW   Fc receptor-like protein; phagocytosis inducer; rheumatoid arthritis;
KW   immune complex related disease; systemic lupus erythematosus; allergy;
KW   haemolytic anaemia; thrombocytopaenia; anaphylaxis; cancer; lymphoma;
KW   leukaemia; infection; immunomodulator; viral entry inhibitor; therapy;
KW   FcR-II.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /note= "signal peptide"
FT   Protein         19. .263
FT                   /note= "mature FcR-II"
PN   WO9831806-A2.
PD   23-JUL-1998.
PF   20-JAN-1998; U01184.
PR   18-JUN-1997; US-049872.
PR   21-JAN-1997; US-034205.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Gentz RL, Murphy M, Ni J, Olsen HS, Ruben SM;
DR   WPI; 98-414105/35.
DR   N-PSDB; V44825.
PT   Nucleic acid encoding Fc receptor-like polypeptides or their
PT   fragments - and related vectors, transformed cells and antibodies,
PT   useful for treating and diagnosing diseases of the haematopoietic
PT   and immune systems
PS   Claim 23; Fig 2A; 141pp; English.
CC   This sequence is the Fc receptor-like II protein (FcR-II) of the
CC   invention. Cells containing the DNA are used to express the recombinant
CC   protein, and to screen for specific (ant)agonists. The proteins are used
CC   to induce phagocytosis, and their (ant)agonists are used to treat immune
CC   complex related diseases (e.g. rheumatoid arthritis, systemic lupus
CC   erythematosus, haemolytic anaemia, thrombocytopaenia, anaphylaxis,
CC   allergy, colorectal or breast cancer, lymphoma, leukaemia, infection by
CC   intracellular pathogens etc). The antagonists are also useful as
CC   immunomodulators and inhibitors of viral (e.g. human immune deficiency or
CC   dengue viruses) entry into cells. The proteins may also be used to screen
CC   for specific binding agents, i.e. (ant)agonists, for raising antibodies
CC   (Ab), and for identification of particular cells or tissues. The Ab can
CC   be used therapeutically as antagonists; as assay reagents for diagnostic
CC   determination of the levels of expression of the proteins and for
CC   affinity purification of the proteins. The DNA and its fragments are
CC   useful as hybridisation probes or primers for isolating related genes, in
CC   situ hybridisation (chromosome mapping) and diagnostically to measure
CC   mRNA expression.
SQ   Sequence   263 AA;

 W69232  Length: 263  February 11, 2000 15:49  Type: P  Check: 3029  ..

       1  MALVLILQLL TLWPLCHTDI TPSVPPASYH PKPWLGAQPA TVVTPGVNVT

      51  LRCRAPQPAW RFGLFKPGEI APLLFRDVSS ELAEFFLEEV TPAQGGSYRC

     101  CYRRPDWGPG VWSQPSDVLE LLVTEELPRP SLVALPGPVV GPGANVSLRC

     151  AGRLRNMSFV LYREGVAAPL QYRHSAQPWA DFTLLGARAP GTYSCYYHTP

     201  SAPYVLSQRS EVLVISWEDS GSSDYTRGNL VRLGLAGLVL ISLGALVTFD

     251  WRSQNRAPAG IRP

!!AA_SEQUENCE 1.0
ID   W62617 standard; Protein; 212 AA.
AC   W62617;
DT   27-OCT-1998  (first entry)
DE   Rattus norvegicus SOCS1 protein.
KW   SOCS; suppressor of cytokine signalling; PCR primer;
KW   autoimmune disease; diagnosis; cancer; treatment;
KW   cytokine mediated cellular responsiveness; hyperimmunity;
KW   immunosuppression; allergies; hypertension.
OS   Rattus norvegicus.
PN   WO9820023-A1.
PD   14-MAY-1998.
PF   31-OCT-1997; AU0729.
PR   14-FEB-1997; AU-005117.
PR   01-NOV-1996; AU-003384.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI   Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR   WPI; 98-286854/25.
DR   N-PSDB; V38663.
PT   Suppressor of cytokine signalling proteins - useful to treat
PT   disease, injury or abnormality involving cytokine mediated cellular
PT   responsiveness e.g. hyperimmunity, immunosuppression, allergies and
```

PT hypertension
PS Claim 13; Page 119-120; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 212 AA;

W62617 Length: 212 February 11, 2000 15:49 Type: P Check: 5610 ..

1 MVARNQVEAD NAISPASEPR RRPEPSSSSS SSSPAAPARP RPCPVVPAPA

51 PGDTHFRTFR SHSDYRRITR TSALLDACGF YWGPLSVHGA HERLRSEPVG

101 TFLVRDSRQR NCFFALSVKM ASGPTSIRVH FQAGRFHLHG NRETFDCLFE

151 LLEHYVAAPR RMLGAPLRQR RVRPLQELCR QRIVAAVGRE NLARIPLNPV

201 LRDYLSSFPF QI

!!AA_SEQUENCE 1.0
ID W62622 standard; Protein; 130 AA.
AC W62622;
DT 27-OCT-1998 (first entry)
DE Mus musculus SOCS8 protein.
KW SOCS; suppressor of cytokine signalling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Mus musculus.
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR WPI; 98-286854/25.
DR N-PSDB; V38675.
PT Suppressor of cytokine signalling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13; Page 149-150; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 130 AA;

W62622 Length: 130 February 11, 2000 15:49 Type: P Check: 4000 ..

1 MSAILKVGHH CWLPVTSAVN PQRMLRPPPT AVFNCAACCC LWGQMLMNTY

51 RVVQLPEEAK GLVPPEILQK YHGFYSSLFA LVRQPRSLQH LCRCALRSHL

101 EGCLPHALPR LPLPPRMLRF LQLDFEDLLY

!!AA_SEQUENCE 1.0
ID W62623 standard; Protein; 207 AA.
AC W62623;
DT 27-OCT-1998 (first entry)
DE Homo sapiens SOCS11 protein.
KW SOCS; suppressor of cytokine signalling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Homo sapiens.
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR WPI; 98-286854/25.
DR N-PSDB; V38681.
PT Suppressor of cytokine signalling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13; Page 158-159; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 207 AA;

W62623 Length: 207 February 11, 2000 15:49 Type: P Check: 6307 ..

1 LEKCGWYWGP MNWEDAEMKL KGKPDGSFLV RDSSDPRYIL SLSFRSQGIT

51 HHTRMEHYRG TFSLWCHPKF EDRCQSVVEF IKRAIMHSKN GKFLYFLRSR

101 VPGLPPTPVQ LLYPVSRFSN VKSLQHLCRF RIRQLVRIDH IPDLPLPKPL

151 ISYIRKFYYY DPQEEVYLSL KEAQRQFPNR SKRWNPPRSE GLPAGHHQGH

201 LVAKLQL

!!AA_SEQUENCE 1.0
ID W62624 standard; Protein; 134 AA.
AC W62624;
DT 27-OCT-1998 (first entry)
DE Mus musculus SOCS13 protein.
KW SOCS; suppressor of cytokine signalling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Mus musculus.
FH Key Location/Qualifiers
FT Misc_difference 45
FT /note= "undefined amino acid"
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR WPI; 98-286854/25.
DR N-PSDB; V38685.
PT Suppressor of cytokine signalling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13; Page 164; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling

CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 134 AA;

W62624 Length: 134 February 11, 2000 15:49 Type: P Check: 8778 ..

1 GGWDLGRNRL YHDGKNQPSK TYPAFLEPDE TFIVPDSFFV ALDMXDGTLS

51 FIVDGQYMGV AFRGLKGKKL YPVVSAVWGH CEIRMRYLNG LDPEPLPLMD

101 LCRRSVRLAL GKERLGAIPA LPLPASLKAY LLYQ

!!AA_SEQUENCE 1.0
ID W62613 standard; Protein; 212 AA.
AC W62613;
DT 27-OCT-1998 (first entry)
DE Mus musculus SOCS1 protein.
KW SOCS; suppressor of cytokine signalling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Mus musculus.
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR WPI; 98-286854/25.
DR N-PSDB; V38659.
PT Suppressor of cytokine signalling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13; Page 110; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 212 AA;

W62613 Length: 212 February 11, 2000 15:49 Type: P Check: 4527 ..

1 MVARNQVAAD NAISPAAEPR RRSEPSSSSS SSSPAAPVRP RPCPAVPAPA

51 PGDTHFRTFR SHSDYRRITR TSALLDACGF YWGPLSVHGA HERLRAEPVG

101 TFLVRDSRQR NCFFALSVKM ASGPTSIRVH FQAGRFHLDG SRETFDCLFE

151 LLEHYVAAPR RMLGAPLRQR RVRPLQELCR QRIVAAVGRE NLARIPLNPV

201 LRDYLSSFPF QI

!!AA_SEQUENCE 1.0
ID W62614 standard; Protein; 198 AA.
AC W62614;
DT 27-OCT-1998 (first entry)
DE Mus musculus SOCS2 protein.
KW SOCS; suppressor of cytokine signalling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Mus musculus.
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR WPI; 98-286854/25.
DR N-PSDB; V38660.
PT Suppressor of cytokine signalling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13; Page 112; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ Sequence 198 AA;

W62614 Length: 198 February 11, 2000 15:49 Type: P Check: 1060 ..

1 MTLRCLEPSG NGADRTRSQW GTAGLPEEQS PEAARLAKAL RELSQTGWYW

51 GSMTVNEAKE KLKEAPEGTF LIRDSSHSDY LLTISVKTSA GPTNLRIEYQ

101 DGKFRLDSII CVKSKLKQFD SVVHLIDYYV QMCKDKRTGP EAPRNGTVHL

151 YLTKPLYTSA PTLQHFCRLA INKCTGTIWG LPLPTRLKDY LEEYKFQV

!!AA_SEQUENCE 1.0
ID W62615 standard; Protein; 225 AA.
AC W62615;
DT 27-OCT-1998 (first entry)
DE Mus musculus SOCS3 protein.
KW SOCS; suppressor of cytokine signalling; PCR primer;
KW autoimmune disease; diagnosis; cancer; treatment;
KW cytokine mediated cellular responsiveness; hyperimmunity;
KW immunosuppression; allergies; hypertension.
OS Mus musculus.
PN WO9820023-A1.
PD 14-MAY-1998.
PF 31-OCT-1997; AU0729.
PR 14-FEB-1997; AU-005117.
PR 01-NOV-1996; AU-003384.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR WPI; 98-286854/25.
DR N-PSDB; V38661.
PT Suppressor of cytokine signalling proteins - useful to treat
PT disease, injury or abnormality involving cytokine mediated cellular
PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT hypertension
PS Claim 13; Page 114-115; 325pp; English.
CC The sequence is that of a suppressor of cytokine signalling
CC protein (SOCS). SOCS can be used to screen for naturally
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC diseases. Alternatively, specific antibodies can be used to
CC screen for SOCS, which is useful as a knowledge of SOCS levels
CC may be important for the diagnosis of certain cancers. Soluble
CC SOCS polypeptides can be used to treat disease, injury or
CC abnormality involving cytokine mediated cellular responsiveness,
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.

```
SQ   Sequence   225 AA;

 W62615  Length: 225  February 11, 2000 15:49  Type: P  Check: 9828  ..

       1  MVTHSKFPAA GMSRPLDTSL RLKTFSSKSE YQLVVNAVRK LQESGFYWSA

      51  VTGGEANLLL SAEPAGTFLI RDSSDQRHFF TLSVKTQSGT KNLRIQCEGG

     101  SFSLQSDPRS TQPVPRFDCV LKLVHHYMPP PGTPSFSLPP TEPSSEVPEQ

     151  PPAQALPGST PKRAYYIYSG GEKIPLVLSR PLSSNVATLQ HLCRKTVNGH

     201  LDSYEKVTQL PGPIREFLDQ YDAPL

!!AA_SEQUENCE 1.0
ID   W62616 standard; Protein; 211 AA.
AC   W62616;
DT   27-OCT-1998  (first entry)
DE   Homo sapiens SOCS1 protein.
KW   SOCS; suppressor of cytokine signalling; PCR primer;
KW   autoimmune disease; diagnosis; cancer; treatment;
KW   cytokine mediated cellular responsiveness; hyperimmunity;
KW   immunosuppression; allergies; hypertension.
OS   Homo sapiens.
PN   WO9820023-A1.
PD   14-MAY-1998.
PF   31-OCT-1997; AU0729.
PR   14-FEB-1997; AU-005117.
PR   01-NOV-1996; AU-003384.
PA   (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PI   Alexander WS, Hilton DJ, Metcalf D, Nicholson SE,
PI   Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
DR   WPI; 98-286854/25.
DR   N-PSDB; V38662.
PT   Suppressor of cytokine signalling proteins - useful to treat
PT   disease, injury or abnormality involving cytokine mediated cellular
PT   responsiveness e.g. hyperimmunity, immunosuppression, allergies and
PT   hypertension
PS   Claim 13; Page 116-117; 325pp; English.
CC   The sequence is that of a suppressor of cytokine signalling
CC   protein (SOCS). SOCS can be used to screen for naturally
CC   occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
CC   diseases. Alternatively, specific antibodies can be used to
CC   screen for SOCS, which is useful as a knowledge of SOCS levels
CC   may be important for the diagnosis of certain cancers. Soluble
CC   SOCS polypeptides can be used to treat disease, injury or
CC   abnormality involving cytokine mediated cellular responsiveness,
CC   e.g. hyperimmunity, immunosuppression, allergies and hypertension.
SQ   Sequence   211 AA;

 W62616  Length: 211  February 11, 2000 15:49  Type: P  Check: 758  ..

       1  MVAHNQVAAD NAVSTAAEPR RRPEPSSSSS SSPAAPARPR PCPAVPAPAP

      51  GDTHFRTFRS HADYRRITRA SALLDACGFY WGPLSVHGAH ERLRAEPVGT

     101  FLVRDSRQRN CFFALSVKMA SGPTSIRVHF QAGRFHLDGS RESFDCLFEL

     151  LEHYVAAPRR MLGAPLRQRR VRPLQELCRQ RIVATVGREN LARIPLNPVL

     201  RDYLSSFPFQ I

!!AA_SEQUENCE 1.0
ID   W61619 standard; Protein; 250 AA.
AC   W61619;
DT   27-OCT-1998  (first entry)
DE   Clone HTPEF86 of TM4SF superfamily.
KW   Human; receptor; immune disorder; cancers; blood disorder;
KW   juvenile rheumatoid arthritis; Graves disease.
OS   Homo sapiens.
PN   WO9831799-A2.
PD   23-JUL-1998.
PF   21-JAN-1998; U00959.
PR   21-JAN-1997; US-034205.
PR   21-JAN-1997; US-034204.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Gentz RL, Ni J, Rosen CA;
DR   WPI; 98-427559/36.
DR   N-PSDB; V48114.
PT   New isolated poly:nucleotide(s) and encoded receptor poly:peptide(s)
PT   - used to develop products for diagnosing or treating e.g. immune
PT   disorders, cancers, blood disorders or immuno-compromised disease
PT   states
PS   Claim 11; Page 38-39; 79pp; English.
CC   Clone HTPEF86 is a member of the TM4SF receptor superfamily.  The
CC   products generated using the receptor can be used for treating abnormal
CC   conditions related to both an excess of and insufficient amounts of
CC   receptor activity.  They can be used in the treatment of e.g. immune
CC   disorders, cancers, blood disorders, juvenile rheumatoid arthritis,
CC   Graves disease or immunocompromised disease states.  The products can
CC   also be used for detection and diagnosis.
SQ   Sequence   250 AA;

 W61619  Length: 250  February 11, 2000 15:49  Type: P  Check: 1162  ..

       1  MNSMTSAVPV ANSVLVVAPH NGYPVTPGIM SHVPLYPNSQ PQVHLVPGNP

      51  PSLVSNVNGQ PVQKALKEGK TLGAIQIIIG LAHIGLGSIM ATVLVGEYLS

     101  ISFYGGFPFW GGLWFIISGS LSVAAENQPY SYCLLSGSLG LNIVSAICSA

     151  VGVILFITDL SIPHPYAYPD YYPYAWGVNP GMAISGVLLV FCLLEFGIAC

     201  ASSHFGCQLV CCQSSNVSVI YPNIYAANPV ITPEPVTSPP SYSSEIQANK

!!AA_SEQUENCE 1.0
ID   W65460 standard; Protein; 375 AA.
AC   W65460;
DT   09-NOV-1998  (first entry)
DE   Human growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; human.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Modified_site    71
FT                    /note= "N-glycosylated"
FT   Cleavage_site    263. .266
FT                    /note= "RXXR proteolytic cleavage site"
PN   WO9835019-A1.
PD   13-AUG-1998.
PF   06-FEB-1998; U02310.
PR   06-FEB-1997; US-795671.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-447217/38.
PT   Transgenic animal growth differentiation factor-11 is inhibited - by
PT   insertion of transgene, also use of GDF-11 inhibitors for treating
PT   muscular wasting, neuromuscular disease, obesity
PS   Example 3; Page 55-56; 89pp; English.
CC   This is the amino acid sequence of human growth differentiation
CC   factor-8 (GDF-8).  It shows a high degree of sequence homology
CC   to the newly identified human growth differentiation factor-11
CC   (GDF-11, see W65458).  Alignment of the GDF-8 and GDF-11 sequences
CC   reveals potential N-linked glycosylation signals and putative
CC   proteolytic processing sites at analogous positions.  The 2
CC   sequences are related not only in the C-terminal region following
CC   the putative cleavage site (90% amino acid sequence identity) but
CC   also in the pro-region of the molecules (45% amino acid sequence
CC   identity.  Claimed transgenic animals in which GDF-11 production is
CC   reduced produce higher than normal levels of muscle and are useful
CC   in the food industry.  GDF-11 polypeptides, polynucleotides and
CC   antibodies can be used to modulate GDF-11 activity or gene
CC   expression for treatment of cell proliferative disorders involving
CC   muscle, nerve and adipose tissue.
SQ   Sequence   375 AA;

 W65460  Length: 375  February 11, 2000 15:49  Type: P  Check: 1814  ..
```

```
      1  MQKLQLCVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACTWRQNTKS

     51  SRIEAIKIQI LSKLRLETAP NISKDVIRQL LPKAPPLREL IDQYDVQRDD

    101  SSDGSLEDDD YHATTETIIT MPTESDFLMQ VDGKPKCCFF KFSSKIQYNK

    151  VVKAQLWIYL RPVETPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

    201  GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPGPGEDGLN

    251  PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

    301  PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

    351  LYFNGKEQII YGKIPAMVVD RCGCS
```

```
!!AA_SEQUENCE 1.0
ID   W64373 standard; Protein; 120 AA.
AC   W64373;
DT   09-NOV-1998  (first entry)
DE   Mycobacterium tuberculosis antigen RDIF5.
KW   Tuberculosis; infection; diagnosis; antigen; RDIF5.
OS   Mycobacterium tuberculosis strain Erdman.
PN   WO9816645-A2.
PD   23-APR-1998.
PF   07-OCT-1997; U18214.
PR   13-MAR-1997; US-818111.
PR   11-OCT-1996; US-729622.
PA   (CORI-) CORIXA CORP.
PI   Campos-Neto A, Dillon DC, Houghton R, Lodes MJ,
PI   Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
DR   WPI; 98-251292/22.
DR   N-PSDB; V44442.
PT   New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT   to develop products for the detection of M. tuberculosis infection
PT   and diagnosis of tuberculosis
PS   Example 3; Page 203-204; 250pp; English.
CC   This polypeptide comprises an antigenic portion of Mycobacterium
CC   tuberculosis antigen RDIF5.  Its amino acid sequence was deduced
CC   from a DNA molecule (see V44442) isolated from a M. tuberculosis
CC   strain Erdman genomic DNA expression library using rabbit antisera
CC   raised against M. tuberculosis fractionated proteins.  Recombinant
CC   RDIF5 was shown to stimulate cell proliferation and interferon
CC   gamma production in T cells from M. tuberculosis-immune donors.
CC   The invention relates to compositions and methods for diagnosing
CC   tuberculosis.  It provides polypeptides (see W64291-W64379)
CC   comprising antigenic or immunogenic portions of M. tuberculosis
CC   antigens, as well as DNA sequences encoding such polypeptides,
CC   recombinant expression vectors and transformed or transfected host
CC   cells.  Also claimed are methods and diagnostic kits for detecting
CC   M. tuberculosis infection in a patient using these polypeptides,
CC   antibodies or oligonucleotide probes and primers.
SQ   Sequence   120 AA;

 W64373  Length: 120  February 11, 2000 15:49  Type: P  Check: 3575  ..

       1  TPEKSFVDDL DIDSLSMVEI AVQTEDKYGV KIPDEDLAGL RTVGDVVAYI

      51  QKLEEENPEA AQALRAKIES ENPDAARADR CVSPTSQARD ARRPLARSAR

     101  LACRRLPASV PTTRRDPRER

!!AA_SEQUENCE 1.0
ID   W38580 standard; Protein; 174 AA.
AC   W38580;
DT   06-NOV-1998  (first entry)
DE   Streptococcus pneumoniae protein of unknown function.
KW   Streptococcus pneumoniae protein; genetic immunisation; antagonist;
KW   immunological response; inoculation; antibody production; inhibitor;
KW   T cell immune response; antimicrobial compound; bacterial adhesion;
KW   extracellular matrix protein; protein-mediated cell invasion; wound;
KW   pathogenesis.
OS   Streptococcus pneumoniae.
FH   Key             Location/Qualifiers
FT   Misc_difference 1. .174
FT                   /note= "residues designated X are unspecified, and
FT                   represented as Xaa in the specification"
PN   WO9743303-A1.
PD   20-NOV-1997.
PF   14-MAY-1997; U07950.
PR   14-MAY-1996; US-017670.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PI   Black MT, Hodgson JE, Knowles DJC, Nicholas RO,
PI   Stodola RK;
DR   WPI; 98-008793/01.
DR   N-PSDB; T98632.
PT   Novel Streptococcus pneumoniae proteins and related DNA - useful for
PT   diagnosing anti-microbial agents for treatment of bacterial
PT   infections
PS   Claim 12; Pages 351-352; 483pp; English.
CC   This sequence represents a Streptococcus pneumoniae protein of
CC   unknown function, and is encoded by a DNA sequence of the invention.
CC   The DNA sequences were isolated from Streptococcus pneumoniae strain
CC   0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
CC   invention can be used to identify compounds which interact with and
CC   inhibit or activate the activity of the proteins. Antagonists can be
CC   used to treat diseases caused by S. pneumoniae proteins, through genetic
CC   immunisation. They can also be used to induce an immunological response
CC   in a mammal by inoculation with the S. pneumoniae proteins or delivery
CC   of the encoding nucleic acids in a vector adequate to produce antibody
CC   and/or T cell immune responses to protect the animal from disease. The
CC   proteins can also be used to identify antimicrobial compounds which are
CC   capable of inhibiting their bioactivity. In particular the proteins of
CC   the invention can be used to prevent adhesion of bacteria to mammalian
CC   extracellular matrix proteins on in-dwelling devices or in wounds, to
CC   block protein-mediated mammalian cell invasion, and to block the normal
CC   progression of pathogenesis in infections initiated other than by the
CC   implantation of in-dwelling devices or other surgical techniques.
SQ   Sequence   174 AA;

 W38580  Length: 174  February 11, 2000 15:49  Type: P  Check: 5404  ..

       1  MGFNQPLSWL ILHAKGLTES FIKASAQTGA DRITYISCNV RNHGAVDIKL

      51  YQELGYELKK VQPVDLFXQT HHVETVALLS KLDVDKHISV EIELDEMDLT

     101  SAESKATYAQ IKEYVWNKFE LKVSTLYIAQ IKKKCGIELR EHYNKXKKDK

     151  QIIPQCTPEK EEAIMDALRX FKMI

!!AA_SEQUENCE 1.0
ID   W70962 standard; Protein; 212 AA.
AC   W70962;
DT   30-OCT-1998  (first entry)
DE   A STAT function regulatory protein designated SIIS-1.
KW   SIIS-1; STAT-induced inhibitor; STAT function;
KW   JAK/STAT signal transmission system; STAT3; STAT6; inhibit;
KW   tyrosine phosphorylation; gp130; cytokine-regulating protein; CIS;
KW   screen; cytokine regulatory; inhibitory activity.
OS   Mus sp.
PN   WO9830688-A1.
PD   16-JUL-1998.
PF   23-OCT-1997; J03860.
PR   10-JAN-1997; JP-014737.
PA   (KISH/) KISHIMOTO T.
PI   Naka T;
DR   WPI; 98-399137/34.
DR   N-PSDB; V42701.
PT   STAT function regulatory protein - used in screening candidate
PT   substances for cytokine regulatory activity
PS   Claim 2; Pages 39-41; 60pp; Japanese.
CC   The present sequence represents a protein (called SIIS-1, STAT-induced
CC   inhibitor of STAT function 1) which regulates STAT protein function in
CC   the JAK/STAT signal transmission system in mammalian cells. The protein
```

```
CC   is induced by STAT3 or STAT6. It inhibits tyrosine phosphorylation of
CC   STAT3 and of gp130. The SIIS-1 protein sequence contains an SH2 domain
CC   and is related to the cytokine-regulating protein CIS. SIIS-1, or
CC   transformant cells expressing it, may be used to screen candidate
CC   substances for cytokine regulatory or inhibitory activity.
SQ   Sequence   212 AA;

 W70962  Length: 212  February 11, 2000 15:49  Type: P  Check: 4556  ..

       1  MVARNQVAAD NAISPAAEPR RRSEPSSSSS SSSPAAPVRP RPCPGVPAPA

      51  PGDTHFRTFR SHSDYRRITR TSALLDACGF YWGPLSVHGA HERLRAEPVG

     101  TFLVRDSRQP NCFFALSVKM ASGPTSIRVH FQAGRFHLDG NRETFDCLFE

     151  LLEHYVAAPR RMLGAPLRQR RVRPLQELCR QRIVAAVGRE NLARIPLNPV

     201  LRDYLSSFPF QI

!!AA_SEQUENCE 1.0
ID   W77666 standard; Protein; 100 AA.
AC   W77666;
DT   30-OCT-1998  (first entry)
DE   Staphylococcus aureus protein of unknown function.
KW   Staphylococcus aureus protein; immune response induction; eye infection;
KW   antibody production; T-cell immune response; gastrointestinal infection;
KW   respiratory infection; inhibitor; bacterial infection; cardiac infection;
KW   central nervous system; kidney infection; urinary tract infection;
KW   antimicrobial compound identification; broad spectrum antibiotic;
KW   therapy.
OS   Staphylococcus aureus.
FH   Key             Location/Qualifiers
FT   Misc_difference 4
FT                   /note= "unspecified, encoded by CNT"
FT   Misc_difference 8
FT                   /note= "unspecified, encoded by ANA"
FT   Misc_difference 12
FT                   /note= "unspecified, encoded by GCN"
FT   Misc_difference 42
FT                   /note= "unspecified, encoded by GCN"
FT   Misc_difference 45
FT                   /note= "unspecified, encoded by GNA"
PN   EP-841394-A2.
PD   13-MAY-1998.
PF   24-SEP-1997; 307485.
PR   24-SEP-1996; US-027032.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PI   Black MT, Burnham MKR, Hodgson JE, Knowles DJC,
PI   Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M,
PI   Ward JM;
DR   WPI; 98-252940/23.
DR   N-PSDB; V53459.
PT   New nucleic acid sequences from Staphylococcus aureus WCHU29 -
PT   useful in vaccines and for treatment of bacterial infections of e.g.
PT   respiratory tract and central nervous system
PS   Claim 11; Page 317-318; 390pp; English.
CC   This sequence represents a Staphylococcus aureus protein of unknown
CC   function, and is encoded by a DNA sequence of the invention.
CC   The DNA sequences were isolated from Staphylococcus aureus WCHU29
CC   (NCIMB 40771). Host cells containing the DNA sequences are used to
CC   produce polypeptides or fragments. The proteins are used in the treatment
CC   of disease, for inducing an immune response by administering them, to
CC   produce antibody and/or T-cell immune response. Antagonists of the
CC   proteins are used for the inhibition of bacterial polypeptides.
CC   Conditions which may be treated include bacterial infections, especially
CC   respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,
CC   urinary tract, skin, bones and joints. The proteins can also be used to
CC   identify antimicrobial compounds which are broad spectrum antibiotics,
CC   especially useful in the treatment of H. pylori infection.
SQ   Sequence   100 AA;

 W77666  Length: 100  February 11, 2000 15:49  Type: P  Check: 5825  ..

       1  VMPXVFWXPF PXIFIGTALP LAGTVATGAI HFTANEVIPI GXMLXNNGLI

      51  AINLAYQNLD RAFVQDGTNI ESKLSLAATP KLASKGAIRE SIRLAIGANN

!!AA_SEQUENCE 1.0
ID   W38716 standard; Protein; 161 AA.
AC   W38716;
DT   10-NOV-1998  (first entry)
DE   S. pneumoniae aspartyl tRNA synthetase.
KW   Streptococcus pneumoniae protein; genetic immunisation; antagonist;
KW   immunological response; inoculation; antibody production; inhibitor;
KW   T cell immune response; antimicrobial compound; bacterial adhesion;
KW   extracellular matrix protein; protein-mediated cell invasion; wound;
KW   pathogenesis.
OS   Streptococcus pneumoniae.
PN   WO9743303-A1.
PD   20-NOV-1997.
PF   14-MAY-1997; U07950.
PR   14-MAY-1996; US-017670.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PI   Black MT, Hodgson JE, Knowles DJC, Nicholas RO,
PI   Stodola RK;
DR   WPI; 98-008793/01.
DR   N-PSDB; T98758.
PT   Novel Streptococcus pneumoniae proteins and related DNA - useful for
PT   diagnosing anti-microbial agents for treatment of bacterial
PT   infections
PS   Claim 12; Pages 450-451; 483pp; English.
CC   This sequence represents a Streptococcus pneumoniae protein that, based
CC   on homology with a Thermus aquaticus protein, is a aspartyl tRNA
CC   synthetase, and is encoded by a DNA sequence of the invention.
CC   The DNA sequences were isolated from Streptococcus pneumoniae strain
CC   0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
CC   invention can be used to identify compounds which interact with and
CC   inhibit or activate the activity of the proteins. Antagonists can be
CC   used to treat diseases caused by S. pneumoniae proteins, through genetic
CC   immunisation. They can also be used to induce an immunological response
CC   in a mammal by inoculation with the S. pneumoniae proteins or delivery
CC   of the encoding nucleic acids in a vector adequate to produce antibody
CC   and/or T cell immune responses to protect the animal from disease. The
CC   proteins can also be used to identify antimicrobial compounds which are
CC   capable of inhibiting their bioactivity. In particular the proteins of
CC   the invention can be used to prevent adhesion of bacteria to mammalian
CC   extracellular matrix proteins on in-dwelling devices or in wounds, to
CC   block protein-mediated mammalian cell invasion, and to block the normal
CC   progression of pathogenesis in infections initiated other than by the
CC   implantation of in-dwelling devices or other surgical techniques.
SQ   Sequence   161 AA;

 W38716  Length: 161  February 11, 2000 15:49  Type: P  Check: 3773  ..

       1  MKYDDAMALY GSDKPDTRFD MLLQDLTEVV KGVDFKVFSE ALAVKAIVVN

      51  GLEPLTADNY SRKDIDKMTE VAKQYGAKGL AWVKVVDGEL NGPVAKFLTG

     101  IQEELTTALA LEDKDLVLFV ADTLEVANAT LGALRGRIAK ELGLIDNDKF

     151  NFLWVVDWPM F

!!AA_SEQUENCE 1.0
ID   W79302 standard; Protein; 67 AA.
AC   W79302;
DT   24-NOV-1998  (first entry)
DE   A Staphylococcus aureus protein.
KW   Alcaligenes eutrophus; cation efflux system czcD protein; treatment;
KW   prevention; bacterial infection; Helicobacter pylori; vaccine.
OS   Staphylococcus aureus.
PN   WO9823738-A2.
PD   04-JUN-1998.
PF   24-NOV-1997; U22092.
PR   25-NOV-1996; US-031469.
```

```
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PI   Warren RL;
DR   WPI; 98-322718/28.
DR   N-PSDB; V59870.
PT   New nucleic acid from Staphylococcus aureus NCIMB 40771 - useful
PT   for, e.g. diagnosis, prevention and treatment of bacterial
PT   infection(s)
PS   Claim 5; Page 33; 114pp; English.
CC   The present sequence represents Staphylococcus aureus WCHU (NCIMB 40771)
CC   protein that has homology to the Alcaligenes eutrophus cation efflux
CC   system czcD protein. The protein is used to generate antibodies and to
CC   screen for antimicrobials. The products are used to treat or prevent
CC   bacterial infections, particularly where caused by S. aureus but also
CC   against Helicobacter pylori. Particular applications are to treat
CC   subjects before surgery or insertion of an in-dwelling device
CC   (alternatively the device itself is impregnated before placement). The
CC   nucleic acid sequence is used as sources of antisense sequences (for
CC   therapeutic use) or regulatory elements for controlling expression of
CC   bacterial genes, and for antibacterial screening. The protein can be
CC   also used as a vaccine.
SQ   Sequence   67 AA;

 W79302  Length: 67  February 11, 2000 15:49  Type: P  Check: 8790  ..

      1  SHNLNMRGAF LHVIGDLLGS VGAITAAILI WAFGWTIADP IASILVSVII

     51  LKSAWGITKS SINILME

!!AA_SEQUENCE 1.0
ID   W69889 standard; Protein; 376 AA.
AC   W69889;
DT   07-DEC-1998  (first entry)
DE   Rat growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; rat; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Rattus sp.
FH   Key             Location/Qualifiers
FT   Cleavage_site   264. .267
FT   Protein         268. .376
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V45820.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 9; Fig 14d; 125pp; English.
CC   This is the amino acid sequence of rat growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V45820) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   376 AA;

 W69889  Length: 376  February 11, 2000 15:49  Type: P  Check: 1496  ..

      1  MIQKPQMYVY IYLFVLIAAG PVDLNEDSER EANVEKEGLC NACAWRQNTR

     51  YSRIEAIKIQ ILSKLRLETA PNISKDAIRQ LLPRAPPLRE LIDQYDVQRD

    101  DSSDGSLEDD DYHATTETII TMPTESDFLM QADGKPKCCF FKFSSKIQYN

    151  KVVKAQLWIY LRAVKTPTTV FVQILRLIKP MKDGTRYTGI RSLKLDMSPG

    201  TGIWQSIDVK TVLQNWLKQP ESNLGIEIKA LDENGHDLAV TFPGPGEDGL

    251  NPFLEVKVTD TPKRSRRDFG LDCDEHSTES RCCRYPLTVD FEAFGWDWII

    301  APKRYKANYC SGECEFVFLQ KYPHTHLVHQ ANPRGSAGPC CTPTKMSPIN

    351  MLYFNGKEQI IYGKIPAMVV DRCGCS

!!AA_SEQUENCE 1.0
ID   W69891 standard; Protein; 375 AA.
AC   W69891;
DT   07-DEC-1998  (first entry)
DE   Pig growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; pig; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Sus scrofa.
FH   Key             Location/Qualifiers
FT   Cleavage_site   263. .266
FT   Protein         267. .375
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V45822.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 9; Fig 14f; 125pp; English.
CC   This is the amino acid sequence of porcine growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V45822) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
```

```
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   375 AA;

 W69891  Length: 375  February 11, 2000 15:49  Type: P  Check: 1805  ..

       1  MQKLQIYVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACMWRQNTKS

      51  SRLEAIKIQI LSKLRLETAP NISKDAIRQL LPKAPPLREL IDQYDVQRDD

     101  SSDGSLEDDD YHATTETIIT MPTESDLLMQ VEGKPKCCFF KFSSKIQYNK

     151  VVKAQLWIYL RPVKTPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

     201  GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPGPGEDGLN

     251  PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

     301  PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

     351  LYFNGKEQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID   W69892 standard; Protein; 375 AA.
AC   W69892;
DT   07-DEC-1998  (first entry)
DE   Ovine growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; sheep; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Ovis aries.
FH   Key             Location/Qualifiers
FT   Cleavage_site   263. .266
FT   Protein         267. .375
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V45823.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 9; Fig 14f; 125pp; English.
CC   This is the amino acid sequence of sheep growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V45823) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   375 AA;

 W69892  Length: 375  February 11, 2000 15:49  Type: P  Check: 1229  ..

       1  MQKLQIFVYI YLFMLLVAGP VDLNENSEQK ENVEKKGLCN ACLWRQNNKS

      51  SRLEAIKIQI LSKLRLETAP NISKDAIRQL LPKAPPLREL IDQYDVQRDD

     101  SSDGSLEDDD YHVTTETVIT MPTESDLLAE VQEKPKCCFF KFSSKIQHNK

     151  VVKAQLWIYL RPVKTPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGI

     201  GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPEPGEEGLN

     251  PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

     301  PKRYKANYCS GECEFLFLQK YPHTHLVHQA NPKGSAGPCC TPTKMSPINM

     351  LYFNGKEQII YGKIPGMVVD RCGCS

!!AA_SEQUENCE 1.0
ID   W30689 standard; Protein; 376 AA.
AC   W30689;
DT   07-DEC-1998  (first entry)
DE   Murine growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; mouse; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   Modified_site   72. .74
FT                   /note= "Asn is N-glycosylated"
FT   Cleavage_site   264. .267
FT   Protein         268. .376
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V42113.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 3; Fig 5a; 125pp; English.
CC   This is the amino acid sequence of mouse growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V42113) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
```

```
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   376 AA;

 W30689  Length: 376  February 11, 2000 15:49  Type: P  Check: 2293  ..

       1  MMQKLQMYVY IYLFMLIAAG PVDLNEGSER EENVEKEGLC NACAWRQNTR

      51  YSRIEAIKIQ ILSKLRLETA PNISKDAIRQ LLPRAPPLRE LIDQYDVQRD

     101  DSSDGSLEDD DYHATTETII TMPTESDFLM QADGKPKCCF FKFSSKIQYN

     151  KVVKAQLWIY LRPVKTPTTV FVQILRLIKP MKDGTRYTGI RSLKLDMSPG

     201  TGIWQSIDVK TVLQNWLKQP ESNLGIEIKA LDENGHDLAV TFPGPGEDGL

     251  NPFLEVKVTD TPKRSRRDFG LDCDEHSTES RCCRYPLTVD FEAFGWDWII

     301  APKRYKANYC SGECEFVFLQ KYPHTHLVHQ ANPRGSAGPC CTPTKMSPIN

     351  MLYFNGKEQI IYGKIPAMVV DRCGCS

!!AA_SEQUENCE 1.0
ID   W69885 standard; Protein; 375 AA.
AC   W69885;
DT   07-DEC-1998  (first entry)
DE   Human growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; human; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   Modified_site   71. .73
FT                   /note= "Asn is N-glycosylated"
FT   Cleavage_site   263. .266
FT   Protein         267. .375
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V45813.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 3; Fig 5c; 125pp; English.
CC   This is the amino acid sequence of human growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V45810) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   375 AA;

 W69885  Length: 375  February 11, 2000 15:49  Type: P  Check: 1814  ..

       1  MQKLQLCVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACTWRQNTKS

      51  SRIEAIKIQI LSKLRLETAP NISKDVIRQL LPKAPPLREL IDQYDVQRDD

     101  SSDGSLEDDD YHATTETIIT MPTESDFLMQ VDGKPKCCFF KFSSKIQYNK

     151  VVKAQLWIYL RPVETPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

     201  GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPGPGEDGLN

     251  PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

     301  PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

     351  LYFNGKEQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID   W69886 standard; Protein; 375 AA.
AC   W69886;
DT   07-DEC-1998  (first entry)
DE   Baboon growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; baboon; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Papio sp.
FH   Key             Location/Qualifiers
FT   Cleavage_site   263. .266
FT   Protein         267. .375
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V45817.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 9; Fig 14a; 125pp; English.
CC   This is the amino acid sequence of baboon growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
```

```
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V45817) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   375 AA;

 W69886  Length: 375  February 11, 2000 15:49  Type: P  Check: 1463  ..

       1  MQKLQLCVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACTWRQNTKS

      51  SRIEAIKIQI LSKLRLETAP NISKDAIRQL LPKAPPLREL IDQYDVQRDD

     101  SSDGSLEDDD YHATTETIIT MPTESDFLMQ VDGKPKCCFF KFSSKIQYNK

     151  VVKAQLWIYL RPVETPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

     201  GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPGPGEDGLN

     251  PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EALGWDWIIA

     301  PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

     351  LYFNGKEQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID   W69887 standard; Protein; 375 AA.
AC   W69887;
DT   07-DEC-1998  (first entry)
DE   Bovine growth differentiation factor-8.
KW   Growth differentiation factor-8; GDF-8; human; transgenic animal;
KW   transforming growth factor-beta; muscle; meat; inhibitor; obesity;
KW   neuromuscular disease; muscular dystrophy; cachexia; AIDS; cancer;
KW   therapy.
OS   Bos taurus.
FH   Key             Location/Qualifiers
FT   Cleavage_site   263. .266
FT   Protein         267. .375
FT                   /label= Mat_protein
PN   WO9833887-A1.
PD   06-AUG-1998.
PF   05-FEB-1998; U02479.
PR   23-MAY-1997; US-862445.
PR   05-FEB-1997; US-795071.
PR   28-APR-1997; US-847910.
PA   (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
PI   Lee S, McPherron AC;
DR   WPI; 98-437444/37.
DR   N-PSDB; V45818.
PT   Transgenic animals with gene for growth differentiation factor-8
PT   disrupted - have increased muscle and reduced cholesterol contents,
PT   also use of GDF-8 inhibitors for treating cancer, obesity,
PT   neuromuscular disease
PS   Example 9; Fig 14b; 125pp; English.

CC   This is the amino acid sequence of bovine growth differentiation
CC   factor-8 (GDF-8), a novel member of the transforming growth factor-
CC   beta superfamily that appears to relate to various cell
CC   proliferative disorders, especially those involving muscle, nerve
CC   and adipose tissue.  The sequence was deduced from a cDNA clone
CC   (see V45818) isolated from a skeletal muscle cDNA library.  The
CC   invention provides novel mammalian and avian GDF-8 proteins (see
CC   W69883-92).  A transgenic non-human animal is claimed in which
CC   GDF-8 expression is disrupted or interfered with.  Also claimed
CC   are: (1) chicken or turkey eggs or meat, beef, milk, pork and lamb
CC   from these animals; (2) method for increasing muscle mass in
CC   animals by administering an antibody (Ab) that binds to GDF-8; (3)
CC   inhibiting the action of GDF-8 by treating foetal or adult muscle
CC   or progenitor cells with a GDF-8 inhibitor; (4) isolated nucleic
CC   acid encoding a GDF-8 protein truncated by loss of the C-terminal
CC   active fragment.  The transgenic animals have increased muscle mass
CC   and for poultry reduced cholesterol contents.  Method (3) is used
CC   to treat muscle wasting or neuromuscular diseases, muscular atrophy
CC   and aging, particularly muscular dystrophy, spinal cord or
CC   traumatic injuries, congestive or obstructive lung disease, AIDS
CC   and cachexia.  Method (4) is used to treat cancer of muscle,
CC   connective tissue and bone, or obesity.  Also (not claimed) GDF-8
CC   can be used to maintain myoblasts intended for transplanting or to
CC   improve efficiency of fusion.  Ab can be used to detect and
CC   quantify GDF-8 (particularly in muscle, for diagnosis or monitoring),
CC   also for immunotherapy and in vivo imaging.
SQ   Sequence   375 AA;

 W69887  Length: 375  February 11, 2000 15:49  Type: P  Check: 9305  ..

       1  MQKLQISVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACLWRENTTS

      51  SRLEAIKIQI LSKLRLETAP NISKDAIRQL LPKAPPLLEL IDQFDVQRDA

     101  SSDGSLEDDD YHARTETVIT MPTESDLLTQ VEGKPKCCFF KFSSKIQYNK

     151  LVKAQLWIYL RPVKTPATVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

     201  GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPEPGEDGLT

     251  PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

     301  PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

     351  LYFNGEGQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID   W80799 standard; Protein; 244 AA.
AC   W80799;
DT   08-DEC-1998  (first entry)
DE   Rhodococcus nitrile hydratase gene fragment product.
KW   Rhodococcus; nitrile hydratase gene; expression vector; microbe;
KW   control factor; nitrilase gene promoter.
OS   Rhodococcus erythropolis.
PN   J10248578-A.
PD   22-SEP-1998.
PF   05-MAR-1997; 065618.
PR   05-MAR-1997; JP-065618.
PA   (NITT ) NITTO CHEM IND CO LTD.
DR   WPI; 98-560730/48.
DR   N-PSDB; V65530.
PT   Expression vector - for Rhodococcus genus microbe
PS   Claim 2; Pages 5-6; 10pp; Japanese.
CC   This represents a product of the nitrile hydratase gene fragment used in
CC   the construction of an expression vector for a Rhodococcus genus microbe.
CC   The expression vector contains a DNA region encoding a control factor
CC   having the ability of activating a nitrilase gene promoter, a DNA region
CC   of the nitrilase gene promotor activated by the above control factor, a
CC   DNA region producible in the cell of a Rhodococcus genus microbe, and a
CC   chemically resistant DNA region functioning in the Rhodococcus genus
CC   microbe. An exotic gene can be expressed constitutionally.
SQ   Sequence   244 AA;
```

```
W80799  Length: 244  February 11, 2000 15:49  Type: P  Check: 2018  ..

       1  MAGADVHAQG GTNRRARILV VDDEKHVRTM VTWQLESENF DVVAAADGDA

      51  ALRQVTESAP DLMVLDLSLP GKGGLEVLAT VRRTDALPIV VLTARRDETE

     101  RIVALDLGAD DYVIKPFSPR ELAARIRAVL RRTTAEPPHE AAVQRFGDLE

     151  IDTAAREVRL HGIPLEFTTK EFDLLAYMAA SPMQVFSRRR LLLEVWRSSP

     201  DWQQDATVTE HVHRIRRKIE EDPTKPTILQ TVRGAGYRFD GERA

!!AA_SEQUENCE 1.0
ID   W80624 standard; Protein; 189 AA.
AC   W80624;
DT   24-DEC-1998  (first entry)
DE   S. pneumoniae GTP binding protein.
KW   Streptococcus pneumoniae protein; recombinant; gene expression; DNA chip;
KW   virulence; antibody; infection; detection; treatment; hypothetical;
KW   cell wall biosynthetic, external target; minimal gene set protein.
OS   Streptococcus pneumoniae.
PN   WO9826072-A1.
PD   18-JUN-1998.
PF   09-DEC-1997; U22578.
PR   13-DEC-1996; US-036281.
PA   (ELIL ) LILLY & CO ELI.
PI   Baltz RH, Burgett SG, Dehoff BS, Hoskins JA, Jaskunas SR,
PI   Mills BJ, Norris FH, Peery RB, Rockey PK, Rosteck PR,
PI   Skatrud PL, Smith MC, Solenberg PJ, Treadway PJ,
PI   Young Bellido ML;
DR   WPI; 98-348529/30.
DR   N-PSDB; V65202.
PT   Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
PT   for evaluating gene expression, and identification of virulence
PT   genes
PS   Claim 3; Pages 210-211; 333pp; English.
CC   This sequence represents a Streptococcus pneumoniae GTP-binding
CC   protein. The invention provides DNA sequences (V65201 to V65304) from
CC   the Streptococcus pneumoniae genome and corresponding protein sequences
CC   (W80605 to W80728). The protein sequences are classified as hypothetical,
CC   cell wall biosynthetic, external target, or minimal gene set proteins. A
CC   recombinant host containing a vector comprising any of the above nucleic
CC   acids can be used for the recombinant expression of the proteins. The
CC   invention also provides a DNA chip having arrayed on it at least 15 base
CC   pair fragment of any one or more of these DNA sequences. The DNA chip can
CC   be used methods for evaluating gene expression in S. pneumoniae and for
CC   identifying virulence genes in S. pneumoniae. Antibodies that selectively
CC   bind to the above proteins or peptide fragments can be used to treat
CC   S. pneumoniae infection. The antibodies can also be used to detect
CC   S. pneumoniae cells.
SQ   Sequence   189 AA;

 W80624  Length: 189  February 11, 2000 15:49  Type: P  Check: 4504  ..

       1  MSASEGRDPY EDYLAINKEL ESYNLRLMER PQIIVTNKMD MPESQENLEE

      51  FKKKLAENYD EFEELPAIFP ISGLTKQGLA TLLDATAELL DKTPEFLLYD

     101  ESDMEEEVYY GFDEEEKAFE ISRDDDATWV LSGEKLMKLF NMTNFDRDES

     151  VMKFARQLRG MGVDEALRAR GAKDGDLVRI GKFEFEFVD

 !!AA_SEQUENCE 1.0
 ID   W68494 standard; Protein; 368 AA.
 AC   W68494;
 DT   07-JAN-1999  (first entry)
 DE   E2 papillomavirus protein encoded by plasmid pCGE2.18VoNCO.
 KW   E2 protein; apoptosis; HPV; infection; papillomavirus-associated cancer;
 KW   cervix; virus-infected cell; p53 tumour repressor.
 OS   Papillomavirus.
 PN   WO9832861-A1.
 PD   30-JUL-1998.
 PF   29-JAN-1998; F00169.
 PR   29-JAN-1997; FR-000964.
 PA   (INSP ) INST PASTEUR.
 PA   (UYME-) UNIV MEXICO NACIONAL AUTONOMA.
 PI   Demeret C, Desaintes C, Goyat S, Thierry F, Yaniv M;
 DR   WPI; 98-427957/36.
 DR   N-PSDB; V60833.
 PT   Papilloma virus E2 protein or nucleic acid encoding it - useful for
 PT   treating and preventing cancer of the cervix
 PS   Disclosure; Fig 12; 111pp; French.
 CC   The present sequence is encoded by an insert of plasmid pCGE2.18VoNCO
 CC   encoding an E2 papillomavirus protein. Compounds derived from E2 are able
 CC   to induce apoptosis in cells that have integrated part of the human
 CC   papillomavirus (HPV) genome. The E2 protein and its derivatives, the
 CC   vectors (including those expressing wild-type E2) and the corresponding
 CC   proteins or truncated E2 (E2TR) are all useful for treating or preventing
 CC   papillomavirus infection, particularly papillomavirus-associated cancers
 CC   (especially of the cervix uteri). The proteins, and sequences expressing
 CC   them, also induce apoptosis of virus-infected cells and increasing the
 CC   activity of the p53 tumour repressor.
 SQ   Sequence   368 AA;

  W68494  Length: 368  February 11, 2000 15:49  Type: P  Check: 9982  ..

        1  TMETPKETLS ERLSCVQDKI IDHYENDSKD IDSQIQYWQL IRWENAIFFA

       51  AREHGIQTLN HQVVPAYNIS KSKAHKAIEL QMALQGLAQS RYKTEDWTLQ

      101  DTCEELWNTE PTHCFKKGGQ TVQVYFDGNK DNCMTYVAWD SVYYMTDAGT

      151  WDKTATCVSH RGLYYVKEGY NTFYIEFKSE CEKYGNTGTW EVHFGNNVID

      201  CNDSMCSTSD DTVSATQLVK QLQHTPSPYS STVSVGTAKT YGQTSAATRP

      251  GHCGLAEKQH CGPVNPLLGA ATPTGNNKRR KLCSGNTTPI IHLKGDRNSL

      301  KCLRYRLRKH SDHYRDISST WHWTGAGNEK TGILTVTYHS ETQRTKFLNT

      351  VAIPDSVQIL VGYMTMGS

 !!AA_SEQUENCE 1.0
 ID   W82504 standard; Protein; 225 AA.
 AC   W82504;
 DT   01-FEB-1999  (first entry)
 DE   Human EPRG1 protein #1.
 KW   EPRG1; EPO primary response gene 1; diagnosis; gene therapy; immunity;
 KW   disease; vaccine; inoculate; antibody; T cell; anaemia; polycythemia;
 KW   cancer; neutropenia; AIDS; diabetes, myelosuppression; allergy; asthma;
 KW   autoimmune disease; inflammatory disease; chromosome mapping; human.
 OS   Homo sapiens.
 PN   EP-877030-A2.
 PD   11-NOV-1998.
 PF   07-MAY-1998; 303597.
 PR   01-MAY-1998; US-071342.
 PR   07-MAY-1997; US-045890.
 PA   (SMIK ) SMITHKLINE BECKMAN CORP.
 PI   Dillon S, Lord K;
 DR   WPI; 98-570499/49.
 DR   N-PSDB; V69307.
 PT   New EPO primary response gene polypeptides and polynucleotides -
 PT   useful as diagnostic reagents and for prevention and treatment of
 PT   cancer and autoimmune and inflammatory diseases
 PS   Disclosure; Page 19-20; 25pp; English.
 CC   This sequence represents a novel human EPO primary response gene 1
 CC   (EPRG1) polypeptide. EPRG1 polypeptides and polynucleotides are useful
 CC   for diagnosing a disease or susceptibility to a disease by detecting
 CC   mutations in the EPRG1 gene using probes containing the EPRG1 nucleotide
 CC   sequence, or determining EPRG1 polypeptide or mRNA expression levels
 CC   EPRG1 polypeptides can be used to screen for agonists and antagonists
 CC   which bind the EPRG1 polypeptide by measuring resulting mRNA levels with
 CC   ELISA. These can be used in treatment to activate (agonist) or inhibit
 CC   (antagonist eg EPRG1 ligand, receptor or substrate) EPRG1 activity, in
 CC   addition to direct administration of antisense sequences to prevent
 CC   expression, or EPRG1 polypeptides to treat conditions associated with
```

CC a lack of EPRG1 protein. Gene therapy may also be used to affect
CC endogenous EPRG1 polypeptide production. EPRG1 antibodies are useful for
CC inducing an immune response to immunise and prevent diseases, and for
CC isolating EPRG1 clones or purifying the polypeptides by affinity
CC chromatography. EPRG1 polypeptides can be administered directly or as a
CC vaccine to inoculate against disease by inducing an antibody and T-cell
CC response. Diseases diagnosed, prevented or treated include anaemia,
CC polycythemia, cancer, neutropenia, AIDS, drug-induced anaemia, diabetes,
CC myelosuppression, autoimmune diseases, rheumatoid arthritis and multiple
CC sclerosis, and inflammatory diseases, including asthma and allergies. The
CC EPRG1 polypeptide is also useful for mapping the gene to a chromosome,
CC allowing gene inheritance to be studied through linkage analysis. The
CC 3'-UTR segment of EPRG1 RNA may be useful to screen for agents which
CC modulate RNA stability and turnover rate.
SQ Sequence 225 AA;

W82504 Length: 225 February 11, 2000 15:49 Type: P Check: 447 ..

1 MVTHSKFPAA GMSRPLDTSL RLKTFSSKSE YQLVVNAVRK LQESGFYWSA

51 VTGGEANLLL SAEPAGTFLI RDSSDQRHFF TLSVKTQSGT KNLRIQCEGG

101 SFSLQSDPRS TQPVPRFDCV LKLVHHYMPP PGAPSFPSPP TEPSSEVPEQ

151 PSAQPLPGSP PRRAYYIYSG GEKIPLVLSR PLSSNVATLQ HLCRKTVNGH

201 LDSYEKVTQL PGPIREFLDQ YDAPL

!!AA_SEQUENCE 1.0
ID W82505 standard; Protein; 157 AA.
AC W82505;
DT 01-FEB-1999 (first entry)
DE Human EPRG1 protein #2.
KW EPRG1; EPO primary response gene 1; diagnosis; gene therapy; immunity;
KW disease; vaccine; inoculate; antibody; T cell; anaemia; polycythemia;
KW cancer; neutropenia; AIDS; diabetes, myelosuppression; allergy; asthma;
KW autoimmune disease; inflammatory disease; chromosome mapping; human.
OS Homo sapiens.
PN EP-877030-A2.
PD 11-NOV-1998.
PF 07-MAY-1998; 303597.
PR 01-MAY-1998; US-071342.
PR 07-MAY-1997; US-045890.
PA (SMIK) SMITHKLINE BECKMAN CORP.
PI Dillon S, Lord K;
DR WPI; 98-570499/49.
DR N-PSDB; V69308.
PT New EPO primary response gene polypeptides and polynucleotides -
PT useful as diagnostic reagents and for prevention and treatment of
PT cancer and autoimmune and inflammatory diseases
PS Claim 1; Page 2; 25pp; English.
CC This sequence represents a novel human EPO primary response gene 1
CC (EPRG1) polypeptide. EPRG1 polypeptides and polynucleotides are useful
CC for diagnosing a disease or susceptibility to a disease by detecting
CC mutations in the EPRG1 gene using probes containing the EPRG1 nucleotide
CC sequence, or determining EPRG1 polypeptide or mRNA expression levels
CC EPRG1 polypeptides can be used to screen for agonists and antagonists
CC which bind the EPRG1 polypeptide by measuring resulting mRNA levels with
CC ELISA. These can be used in treatment to activate (agonist) or inhibit
CC (antagonist eg EPRG1 ligand, receptor or substrate) EPRG1 activity, in
CC addition to direct administration of antisense sequences to prevent
CC expression, or EPRG1 polypeptides to treat conditions associated with
CC a lack of EPRG1 protein. Gene therapy may also be used to affect
CC endogenous EPRG1 polypeptide production. EPRG1 antibodies are useful for
CC inducing an immune response to immunise and prevent diseases, and for
CC isolating EPRG1 clones or purifying the polypeptides by affinity
CC chromatography. EPRG1 polypeptides can be administered directly or as a
CC vaccine to inoculate against disease by inducing an antibody and T-cell
CC response. Diseases diagnosed, prevented or treated include anaemia,
CC polycythemia, cancer, neutropenia, AIDS, drug-induced anaemia, diabetes,
CC myelosuppression, autoimmune diseases, rheumatoid arthritis and multiple
CC sclerosis, and inflammatory diseases, including asthma and allergies. The
CC EPRG1 polypeptide is also useful for mapping the gene to a chromosome,
CC allowing gene inheritance to be studied through linkage analysis. The
CC 3'-UTR segment of EPRG1 RNA may be useful to screen for agents which
CC modulate RNA stability and turnover rate.
SQ Sequence 157 AA;

W82505 Length: 157 February 11, 2000 15:49 Type: P Check: 9756 ..

1 EFANSSDQRH FFQLSVKTQS GTKNLRIQCE GGSFCLQSDP RKTQPVPRVD

51 CVLKLVHRSG APAGAPSFPS PPTEPSSEVP EQPSAQPLPG SPPRRAYYIY

101 SGGEKIPLVL SRPLSSNVAT LQHLCRKTVN GHLDSYEKVT QLPGPIREFL

151 DQYDAPL

!!AA_SEQUENCE 1.0
ID W81740 standard; Protein; 120 AA.
AC W81740;
DT 27-JAN-1999 (first entry)
DE M. tuberculosis immunogenic polypeptide RDIF5.
KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW vaccine; pharmaceutical; infection; diagnosis.
OS Mycobacterium tuberculosis.
PN WO9816646-A2.
PD 23-APR-1998.
PF 07-OCT-1997; U18293.
PR 13-MAR-1997; US-818112.
PR 11-OCT-1996; US-730510.
PA (CORI-) CORIXA CORP.
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ,
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
DR WPI; 98-261042/23.
DR N-PSDB; V64551.
PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and for diagnosis, treatment and prevention of tuberculosis
PS Example 3d; Page 190-191; 230pp; English.
CC This sequence represents an immunogenic portion of a soluble
CC Mycobacterium tuberculosis (MT) antigen which can be used in a method
CC for inducing protective immunity against tuberculosis (TB). This sequence
CC can be formulated into vaccines and/or pharmaceutical compositions for
CC immunising against M. tuberculosis infection or may be used for the
CC diagnosis of tuberculosis.
SQ Sequence 120 AA;

W81740 Length: 120 February 11, 2000 15:49 Type: P Check: 3575 ..

1 TPEKSFVDDL DIDSLSMVEI AVQTEDKYGV KIPDEDLAGL RTVGDVVAYI

51 QKLEEENPEA AQALRAKIES ENPDAARADR CVSPTSQARD ARRPLARSAR

101 LACRRLPASV PTTRRDPRER

!!AA_SEQUENCE 1.0
ID W72902 standard; Protein; 226 AA.
AC W72902;
DT 21-JAN-1999 (first entry)
DE Mycobacterium tuberculosis antigen CFP19A.
KW Mycobacterium tuberculosis; antigen; vaccine; immunological;
KW immunogen; infection.
OS Mycobacterium tuberculosis.
FH Key Location/Qualifiers
FT Peptide 1. .38
FT /label= signal
PN WO9844119-A1.
PD 08-OCT-1998.
PF 01-APR-1998; DK0132.
PR 05-JAN-1998; US-070488.
PR 02-APR-1997; DK-000376.
PR 18-APR-1997; US-044624.
PR 10-NOV-1997; DK-001277.
PA (STAT-) STATENS SERUM INST.
PI Andersen P, Florio W, Nielsen R, Oettinger T, Rasmussen PB,

PI Rosenkrands I, Weldingh K;
DR WPI; 98-542705/46.
DR N-PSDB; V63926.
PT New isolated mycobacteria polypeptides and nucleic acids - used for
PT developing products for the diagnosis of or vaccination against
PT mycobacterial infections, particularly tuberculosis
PS Claim 1; Page 159; 163pp; English.
CC The present sequence represents a Mycobacterium tuberculosis protein.
CC Products from the present invention, which describes protein fragments
CC and nucleic acid fragments derived from M.tuberculosis, can be used in
CC the detection of and prevention of mycobacterial infections. In
CC particular, the proteins and nucleic acids can be used for the diagnosis
CC of or vaccination against tuberculosis caused by M. tuberculosis,
CC M. africanum or M. bovis.
SQ Sequence 226 AA;

W72902 Length: 226 February 11, 2000 15:49 Type: P Check: 5027 ..

1 MIPRPQPHSG RWRAGAARRL TSLVAAAFAA ATLLLTPALA PPASAGCPDA

51 EVVFARGTGE PPGLGRVGQA FVSSLRQQTN KSIGTYGVNY PANGDFLAAA

101 DGANDASDHI QQMASACRAT RLVLGGYSQG AAVIDIVTAA PLPGLGFTQP

151 LPPAADDHIA AIALFGNPSG RAGGLMSALT PQFGSKTINL CNNGDPICSD

201 GNRWRAHLGY VPGMTNQAAR FVASRI

!!AA_SEQUENCE 1.0
ID W80495 standard; Protein; 207 AA.
AC W80495;
DT 29-JAN-1999 (first entry)
DE Human vascular endothelial growth factor (VEGF)-B186.
KW Vascular endothelial growth factor; VEGF; proliferation; human;
KW endothelial cell; angiogenesis; tissue growth; organ repair.
OS Homo sapiens.
PN US5840693-A.
PD 24-NOV-1998.
PF 01-MAR-1996; 609443.
PR 01-MAR-1996; US-609443.
PR 01-MAR-1995; US-397651.
PR 06-JUN-1995; US-469427.
PR 06-DEC-1995; US-569063.
PA (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
PI Alitalo K, Eriksson U, Olofsson B, Pajusola K;
DR WPI; 99-034079/03.
DR N-PSDB; V63569.
PT Vascular endothelial growth factor-B iso:forms, and DNA encoding
PT them - useful for inducing angiogenesis and cellular proliferation,
PT and raising antibodies to inhibit activities in e.g. tumours
PS Claim 1; Fig 15; 52pp; English.
CC The present sequence represents human vascular endothelial growth factor
CC (VEGF)-B186. VEGF proteins are used for promoting proliferation of
CC endothelial cells and for stimulating angiogenesis (the proliferation
CC of new capillaries form pre-existing blood vessels). These activities
CC are useful for treating tissue growth and repair, including organ
CC repair. This is also useful in pregnancy, in follicle development,
CC as these processes must occur in development of the placenta. The
CC proteins can also be used to raise antibodies, either for use in
CC detection of the proteins or as inhibitors of their action. This is
CC especially useful as angiogenesis is required by tumours as they need
CC new blood supplies to grow and proliferate.
SQ Sequence 207 AA;

W80495 Length: 207 February 11, 2000 15:49 Type: P Check: 1679 ..

1 MSPLLRRLLL AALLQLAPAQ APVSQPDAPG HQRKVVSWID VYTRATCQPR

51 EVVVPLTVEL MGTVAKQLVP SCVTVQRCGG CCPDDGLECV PTGQHQVRMQ

101 ILMIRYPSSQ LGEMSLEEHS QCECRPKKKD SAVKPDRAAT PHHRPQPRSV

151 PGWDSAPGAP SPADITHPTP APGPSAHAAP STTSALTPGP AAAAADAAAS

201 SVAKGGA

!!AA_SEQUENCE 1.0
ID W85306 standard; peptide; 24 AA.
AC W85306;
DT 16-FEB-1999 (first entry)
DE Helper T-cell class II peptide derived from mouse invariant chain.
KW Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1;
KW DR7; cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW acquired immune deficiency syndrome; malaria; cancer;
KW allograft rejection; allergy; Lyme disease; hepatitis;
KW post-streptococcal endocarditis; glomerulonephritis;
KW food hypersensitivity.
OS Synthetic.
OS Mus sp.
PN WO9832456-A1.
PD 30-JUL-1998.
PF 23-JAN-1998; U01373.
PR 07-FEB-1997; US-037432.
PR 23-JAN-1997; US-036713.
PA (EPIM-) EPIMMUNE INC.
PI Sette A, Sidney J, Southwood S;
DR WPI; 98-427679/36.
PT Composition containing peptide that induces cytotoxic T lymphocyte
PT response, and helper peptide - can bind to human leucocyte antigen
PT alleles, used to treat or prevent cancers, parasitic infections and
PT autoimmune disease
PS Disclosure; Page 40; 51pp; English.
CC W85284-451 represent helper T-cell class II peptides, which can bind to
CC the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides
CC are used in the course of the invention. The specification describes
CC peptides that that induce a cytotoxic T lymphocyte (CTL) response, and
CC T-helper peptides, that are used together to generate a CTL response for
CC the treatment or prevention of viral, fungal, bacterial or parasitic
CC infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC cancer or condyloma acuminatum). Helper T-cell peptides may be used
CC alone to induce a helper T cell response, e.g. in cases of autoimmune
CC disease, allograft rejection, allergy, Lyme disease, hepatitis,
CC post-streptococcal endocarditis, glomerulonephritis and food
CC hypersensitivity.
SQ Sequence 24 AA;

W85306 Length: 24 February 11, 2000 15:49 Type: P Check: 3126 ..

1 LPKPPKPVSK MRMATPLLMQ ALPM

!!AA_SEQUENCE 1.0
ID W86202 standard; protein; 206 AA.
AC W86202;
DT 16-FEB-1999 (first entry)
DE Human VEGF-related factor (VRF)-2 sequence.
KW VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN WO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998; U07801.
PR 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI; 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Disclosure; Fig 1; 113pp; English.
CC This represents the amino acid sequence of human VRF-2 (a vascular
CC endothelial growth factor (VEGF)-related factor). The invention

CC provides truncated VRP (VEGF-related protein) subunits that have at
CC least one amino acid N-terminal to the first Cys of the core sequence
CC deleted. Host cells transformed or transfected with expression vectors
CC containing nucleic acids encoding the truncated VRP subunits are used to
CC produce the truncated proteins recombinantly. The truncated VRP subunits,
CC optionally expressed from gene therapy vectors, have in vivo and in vitro
CC angiogenic activity and are used to stimulate angiogenesis, particularly
CC coronary collateral vessel development in cases of cardiac ischaemia; to
CC stimulate endothelial cell growth and migration in vitro; to treat heart
CC disease; to treat ischaemia (e.g.cardiac, chronic coronary or chronic
CC lower limb ischaemia; stroke and peripheral vascular disease); to promote
CC healing of wounds (of skin or intestines), and to increase vascular
CC permeability.
SQ Sequence 206 AA;

W86202 Length: 206 February 11, 2000 15:49 Type: P Check: 9273 ..

1 MSPLLRRLLL AALLQLAPAQ APVSQPDAPG HQRKVVSWID VYTRATCQPR

51 EVVVPLTVEL MGTVAKQLVP SCVTVQRCGG CCPDDGLECV PTGQHQVRMQ

101 ILMIRYPSSQ LGEMSLEEHS QCECRPKKDS AVKPDRAATP HHRPQPRSVP

151 GWDSAPGAPS PADITHPTPA PGPSAHAAPS TTSALTPGPA AAAADAAASS

201 VAKGGA

!!AA_SEQUENCE 1.0
ID W68541 standard; Protein; 420 AA.
AC W68541;
DT 12-FEB-1999 (first entry)
DE Amino acid sequence of the SIAX DP2-64 (Oct-T1) gene.
KW Tumour rejection antigen precursor; TRAP; TRH; leukaemia; screening;
KW lymphoma; cancer; HLA; human lymphocyte antigen; vaccine;
KW SIAX DP2-64; Oct-T1.
OS Homo sapiens.
PN WO9849299-A1.
PD 05-NOV-1998.
PF 22-APR-1998; U07784.
PR 25-APR-1997; US-845998.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Boon-Falleur T, Coulie PG, De Smet C, Lucas S, Van Baren N;
DR WPI; 99-009425/01.
DR N-PSDB; V33912.
PT New diagnosis of leukaemia - by detecting genes for tumour antigen
PT rejection precursors or corresponding proteins
PS Claim 11; Pages 68-69; 88pp; English.
CC The present sequence represents the amino acid sequence of the
CC SIAX DP2-64 (Oct-T1) gene. The protein is a tumour rejection
CC antigen precursor (TRAP). The specification describes the treatment of
CC disorders which characterised by expression of a leukaemia-associated
CC nucleic acid such as TRH. The products are used for in vivo or in vitro
CC screening for leukaemia, lymphoma or other cancers by usual
CC hybridisation/amplification or immunoassay methods. TRAPs, when
CC processed to antigens or complexed with HLA (human lymphocyte antigen)
CC molecules, or nucleic acid encoding them, are useful in vaccines for
CC treating leukaemia.
SQ Sequence 420 AA;

W68541 Length: 420 February 11, 2000 15:49 Type: P Check: 8848 ..

1 MMSMNSKQPH FAMHPTLPEH KYPSLHSSSE AIRRACLPTP PLQSNLFASL

51 DETLLARAEA LAAVDIAVSQ GKSHPFKPDA TYHTMNSVPC TSTSTVPLAH

101 HHHHHHHHQA LEPGDLLDHI SSPSLALMAG AGGAGAAAGG GGAHDGPGGG

151 GGPGGGGGPG GGGPGGGGGG GPGGGGGGPG GGLLGGSAHP HPHMHSLGHL

201 SHPAAAAAMN MPSGLPHPGL VAAAAHHGAA AAAAAAAAGQ VAAASAAAAV

251 VGAAGLASIC DSDTDPRELE AFAERFKQRR IKLGVTQADV GSALANLKIP

301 GVGSLSQSTI CRFESLTLSH NNMIALKPIL QAWLEEAEGA QREKMNKPEL

351 FNGGEKKRKR TSIAAPEKRS LEAYFAVQPR PSSEKIAAIA EKLDLKKNVV

401 RVWFCNQRQK QKRMKFSATY

!!AA_SEQUENCE 1.0
ID W86235 standard; protein; 185 AA.
AC W86235;
DT 16-FEB-1999 (first entry)
DE Human VRF (VEGF-related factor)-2 full length sequence.
KW VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; vascular permeability; VRF.
OS Homo sapiens.
PN WO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998; U07801.
PR 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI; 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Claim 5; Fig 2B; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischaemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischaemia (e.g.cardiac, chronic coronary or chronic lower limb
CC ischaemia; stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W86234 to W86239 represent full length VRP sequences from
CC which the truncated fragments are created.
SQ Sequence 185 AA;

W86235 Length: 185 February 11, 2000 15:49 Type: P Check: 8227 ..

1 PVSQPDAPGH QRKVVSWIDV YTRATCQPRE VVVPLTVELM GTVAKQLVPS

51 CVTVQRCGGC CPDDGLECVP TGQHQVRMQI LMIRYPSSQL GEMSLEEHSQ

101 CECRPKKDSA VKPDRAATPH HRPQPRSVPG WDSAPGAPSP ADITHPTPAP

151 GPSAHAAPST TSALTPGPAA AAADAAASSV AKGGA

!!AA_SEQUENCE 1.0
ID W86214 standard; protein; 178 AA.
AC W86214;
DT 16-FEB-1999 (first entry)
DE Human VRF-2 truncated fragment 1.
KW VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN WO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998; U07801.
PR 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI; 99-009426/01.

PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Claim 5; Fig 2B; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischaemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischaemia (e.g.cardiac, chronic coronary or chronic lower limb
CC ischaemia; stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
CC related factor).
SQ Sequence 178 AA;

W86214 Length: 178 February 11, 2000 15:49 Type: P Check: 5797 ..

1 PGHQRKVVSW IDVYTRATCQ PREVVVPLTV ELMGTVAKQL VPSCVTVQRC

51 GGCCPDDGLE CVPTGQHQVR MQILMIRYPS SQLGEMSLEE HSQCECRPKK

101 DSAVKPDRAA TPHHRPQPRS VPGWDSAPGA PSPADITHPT PAPGPSAHAA

151 PSTTSALTPG PAAAAADAAA SSVAKGGA

!!AA_SEQUENCE 1.0
ID W86215 standard; protein; 173 AA.
AC W86215;
DT 16-FEB-1999 (first entry)
DE Human VRF-2 truncated fragment 2.
KW VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN WO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998; U07801.
PR 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI; 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Claim 5; Fig 2B; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischaemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischaemia (e.g.cardiac, chronic coronary or chronic lower limb
CC ischaemia; stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
CC related factor).
SQ Sequence 173 AA;

W86215 Length: 173 February 11, 2000 15:49 Type: P Check: 4363 ..

1 KVVSWIDVYT RATCQPREVV VPLTVELMGT VAKQLVPSCV TVQRCGGCCP

51 DDGLECVPTG QHQVRMQILM IRYPSSQLGE MSLEEHSQCE CRPKKDSAVK

101 PDRAATPHHR PQPRSVPGWD SAPGAPSPAD ITHPTPAPGP SAHAAPSTTS

151 ALTPGPAAAA ADAAASSVAK GGA

!!AA_SEQUENCE 1.0
ID W86216 standard; protein; 168 AA.
AC W86216;
DT 16-FEB-1999 (first entry)
DE Human VRF-2 truncated fragment 3.
KW VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN WO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998; U07801.
PR 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;
DR WPI; 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Claim 5; Fig 2B; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischaemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischaemia (e.g.cardiac, chronic coronary or chronic lower limb
CC ischaemia; stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
CC related factor).
SQ Sequence 168 AA;

W86216 Length: 168 February 11, 2000 15:49 Type: P Check: 2203 ..

1 IDVYTRATCQ PREVVVPLTV ELMGTVAKQL VPSCVTVQRC GGCCPDDGLE

51 CVPTGQHQVR MQILMIRYPS SQLGEMSLEE HSQCECRPKK DSAVKPDRAA

101 TPHHRPQPRS VPGWDSAPGA PSPADITHPT PAPGPSAHAA PSTTSALTPG

151 PAAAAADAAA SSVAKGGA

!!AA_SEQUENCE 1.0
ID W86217 standard; protein; 163 AA.
AC W86217;
DT 16-FEB-1999 (first entry)
DE Human VRF-2 truncated fragment 4.
KW VEGF; VRP; vascular endothelial growth factor; VEGF-related protein;
KW recombinant; truncated; gene therapy; angiogenesis; cardiac ischaemia;
KW coronary; collateral vessel development; cell growth; migration; heart;
KW lower limb ischaemia; stroke; peripheral vascular disease; intestine;
KW wound healing; skin; VEGF-related factor; VRF; vascular permeability.
OS Homo sapiens.
PN WO9849300-A2.
PD 05-NOV-1998.
PF 20-APR-1998; U07801.
PR 25-APR-1997; US-842984.
PA (COLL-) COLLATERAL THERAPEUTICS.
PI Bohlen P;

DR WPI; 99-009426/01.
PT New truncated vascular endothelial growth factor-related protein
PT subunits - lack part of the N-terminal sequence, used to stimulate
PT angiogenesis, e.g. for treating heart disease and ischaemia
PS Claim 5; Fig 2B; 113pp; English.
CC The invention relates to truncated VRP (vascular endothelial growth
CC factor (VEGF)-related protein) subunits that have at least one amino acid
CC N-terminal to the first Cys of the core sequence deleted. Host cells
CC transformed or transfected with expression vectors containing nucleic
CC acids encoding the truncated VRP subunits are used to produce the
CC truncated proteins recombinantly. The truncated VRP subunits, optionally
CC expressed from gene therapy vectors, have in vivo and in vitro angiogenic
CC activity and are used to stimulate angiogenesis, particularly coronary
CC collateral vessel development in cases of cardiac ischaemia; to stimulate
CC endothelial cell growth and migration in vitro; to treat heart disease;
CC to treat ischaemia (e.g.cardiac, chronic coronary or chronic lower limb
CC ischaemia; stroke and peripheral vascular disease); to promote healing of
CC wounds (of skin or intestines), and to increase vascular permeability.
CC Sequences W86214 to W86217 represent truncated fragments of VRF-2 (VEGF-
CC related factor).
SQ Sequence 163 AA;

W86217 Length: 163 February 11, 2000 15:49 Type: P Check: 3123 ..

1 RATCQPREVV VPLTVELMGT VAKQLVPSCV TVQRCGGCCP DDGLECVPTG

51 QHQVRMQILM IRYPSSQLGE MSLEEHSQCE CRPKKDSAVK PDRAATPHHR

101 PQPRSVPGWD SAPGAPSPAD ITHPTPAPGP SAHAAPSTTS ALTPGPAAAA

151 ADAAASSVAK GGA

!!AA_SEQUENCE 1.0
ID W85119 standard; Protein; 446 AA.
AC W85119;
DT 11-FEB-1999 (first entry)
DE A delta-5 desaturase enzyme.
KW Fatty acid; delta-5 desaturase; polyunsaturated fatty acid;
KW malnutrition; inflammation; rheumatoid arthritis; asthma; psoriasis;
KW cancer; diabetes; eczema; platelet aggregation; vasodilation;
KW cholesterol level; endometriosis; premenstrual syndrome;
KW myalgic encephalomyelitis; chronic fatigue; AIDS; multiple sclerosis;
KW acute respiratory syndrome; hypertension; inflammatory skin disorder.
OS Mortierella alpina.
PN WO9846765-A1.
PD 22-OCT-1998.
PF 10-APR-1998; U07422.
PR 11-APR-1997; US-833610.
PA (ABBO) ABBOTT LAB.
PA (CALJ) CALGENE LLC.
PI Chaudhary S, Huang Y, Knutzon D, Leonard AE, Mukerji P,
PI Thurmond J;
DR WPI; 99-009334/01.
DR N-PSDB; V82628.
PT New nucleic acid encoding delta5 and other desaturase enzymes -
PT useful in production of oils of increased arachidonic acid content,
PT used, e.g. for treating cancer, as foods, animal feeds and cosmetics
PS Claim 3; Fig 3A-D; 153pp; English.
CC The present sequence represents a Mortierella alpina fatty acid delta-5
CC desaturase enzyme. The enzyme sequence is used in the methods of
CC the invention. The specification describes methods for desaturating a
CC fatty acid and for producing a desaturated fatty acid by expressing
CC increased levels of a desaturase. The present desaturase is an enzyme
CC which introduces a double bond carbons 5 and 6 from the carboxyl end of
CC a fatty acid molecule. The enzyme can be used for desaturating fatty
CC acids. The enzyme can be used to produce polyunsaturated fatty acids,
CC which can be used for treating malnutrition, in pharmaceutical
CC compositions, in cosmetics or in animal feed. The polyunsaturated fatty
CC acids can be used for treating e.g. restenosis after angioplasty,
CC inflammation, rheumatoid arthritis, asthma, psoriasis, cancer, diabetes
CC or eczema or reduce blood pressure. They can also be used to inhibit
CC platelet aggregation, cause vasodilation, lower cholesterol levels,
CC inhibit proliferation of vessel wall smooth muscle and fibrous tissue,
CC reduce or prevent gastro-intestinal bleeding and other side effects
CC caused by non-steroidal anti-inflammatory drugs, prevent or treat
CC endometriosis and premenstrual syndrome, treat myalgic encephalomyelitis
CC and chronic fatigue after viral infections, treat AIDS, multiple
CC sclerosis, acute respiratory syndrome, hypertension and inflammatory skin
CC disorders.
SQ Sequence 446 AA;

W85119 Length: 446 February 11, 2000 15:49 Type: P Check: 5663 ..

1 MGTDQGKTFT WEELAAHNTK DDLLLAIRGR VYDVTKFLSR HPGGVDTLLL

51 GAGRDVTPVF EMYHAFGAAD AIMKKYYVGT LVSNELPIFP EPTVFHKTIK

101 TRVEGYFTDR NIDPKNRPEI WGRYALIFGS LIASYYAQLF VPFVVERTWL

151 QVVFAIIMGF ACAQVGLNPL HDASHFSVTH NPTVWKILGA THDFFNGASY

201 LVWMYQHMLG HHPYTNIAGA DPDVSTSEPD VRRIKPNQKW FVNHINQHMF

251 VPFLYGLLAF KVRIQDINIL YFVKTNDAIR VNPISTWHTV MFWGGKAFFV

301 WYRLIVPLQY LPLGKVLLLF TVADMVSSYW LALTFQANHV VEEVQWPLPD

351 ENGIIQKDWA AMQVETTQDY AHDSHLWTSI TGSLNYQAVH HLFPNVSQHH

401 YPDILAIIKN TCSEYKVPYL VKDTFWQAFA SHLEHLRVLG LRPKEE

!!AA_SEQUENCE 1.0
ID W73507 standard; Protein; 317 AA.
AC W73507;
DT 01-MAR-1999 (first entry)
DE Human ATG-1709 protein.
KW Human; ATG-1709 protein; secreted ligand; 7-Transmembrane receptor;
KW heart disease; hypertension; cardiovascular disease; kidney disease;
KW obesity; insulin resistance; diabetes; Central Nervous System disorder;
KW therapy; sFRP-1.
OS Homo sapiens.
PN EP-879885-A1.
PD 25-NOV-1998.
PF 16-JAN-1998; 300313.
PR 08-AUG-1997; US-907808.
PR 23-MAY-1997; US-047691.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Hu E, Zhu Y;
DR WPI; 98-596877/51.
DR N-PSDB; V08946.
PT New human secreted protein ATG-1709 polypeptide and polynucleotide -
PT useful as diagnostic reagents and for diagnosing, prevention and
PT treatment of Central Nervous System diseases and diabetes
PS Claim 11; Page 6; 28pp; English.
CC This sequence represents the human ATG-1709 protein of the invention.
CC ATG-1709 is related to human secreted ligands for 7-Transmembrane
CC receptors and similar to murine sFRP-1. ATG-1709 polypeptides and
CC polynucleotides are useful for diagnosing susceptibility to diseases by
CC detecting mutations in the ATG-1709 gene using probes containing the
CC ATG-1709 nucleotide sequence, and can diagnose diseases associated with
CC ATG-1709 imbalance by determining ATG-1709 polypeptide expression levels.
CC ATG-1709 polypeptides can be used to screen for agonists and antagonists
CC which bind the ATG-1709 polypeptide. These can be used in treatment to
CC activate or inhibit ATG-1709 activity, in addition to direct
CC administration of antisense sequences to prevent expression, or ATG-1709
CC polypeptides to treat conditions associated with a lack of ATG-1709.
CC Gene therapy may also be used to affect endogenous ATG-1709 expression.
CC ATG-1709 antibodies are useful for inducing an immune response to
CC immunise and prevent diseases, and for isolating ATG-1709 clones or
CC purifying the polypeptides by affinity chromatography. ATG-1709
CC polypeptides can be administered directly or as a vaccine to inoculate
CC against disease. Diseases diagnosed, prevented or treated include:
CC heart disease; hypertension; cardiovascular diseases; kidney diseases;
CC obesity; insulin resistance; diabetes and Central Nervous System (CNS)
CC diseases. The ATG-1709 polypeptide is also useful for mapping the gene to
CC a chromosome, allowing gene inheritance to be studied through linkage

```
CC   analysis.
SQ   Sequence   317 AA;

 W73507  Length: 317  February 11, 2000 15:49  Type: P  Check: 3945  ..

       1  MRAAAAAGGV RTAALALLLG ALHWAPARCE EYHYYGWQAE PLHGRSYSKP

      51  PQCLDIPADL PLCHTVGYKR MRLPNLLEHE SLAEVKQQAS SWLPLLAKRC

     101  HSDTQVFLCS LFAPVCLDRP IYPCRSLCEA VRAGCAPLME AYGFPWPEML

     151  HCHKFPLDND LCIAVQFGHL PATAPPVTKI CAQCEMEHSA DGLMEQMCSS

     201  DFVVKMRIKE IKIENGDRKL IGAQKKKKLL KPGPLKRKDT KRLVLHMKNG

     251  AGCPCPQLDS LAGSFLVMGR KVDGQLLLMA VYRWDKKNKE MKFAVKFMFS

     301  YPCSLYYPFF YGAAEPH

!!AA_SEQUENCE 1.0
ID   W90011 standard; Protein; 343 AA.
AC   W90011;
DT   18-FEB-1999  (first entry)
DE   Expressed antigen for cluster 35a.
KW   Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW   peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
OS   Helicobacter pylori.
PN   WO9849314-A2.
PD   05-NOV-1998.
PF   27-APR-1998; U08487.
PR   14-OCT-1997; US-061958.
PR   25-APR-1997; US-045107.
PA   (GENE-) GENELABS TECHNOLOGIES INC.
PI   Chow TP, Fry KE, Lim MY, McAtee CP;
DR   WPI; 99-009433/01.
PT   New Helicobacter pylori antigens and related nucleic acid sequences
PT   - useful in serological diagnosis and protective vaccines, providing
PT   long-lasting immune response
PS   Claim 16; Page 338; 402pp; English.
CC   The present sequence represents a Helicobacter pylori antigenic protein
CC   that is characterised by immunoreactivity with H. pylori-positive
CC   antisera. The proteins are highly immunogenic and induce a long-lasting
CC   immune response that persists even after antimicriobial treatment. In
CC   antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC   highly sensitive and specific. The specification also describes 69
CC   previously unrecognised immunogenic cluster families. H. pylori antigens
CC   are used to detect H. pylori-specific antibodies, for diagnosing
CC   infection or to confirm eradication of infection, and in vaccines to
CC   protect against H. pylori infection and related diseases (gastritis,
CC   peptic ulcer, gastric adenocarcinoma/lymphoma).
SQ   Sequence   343 AA;

 W90011  Length: 343  February 11, 2000 15:49  Type: P  Check: 9038  ..

       1  MATKLTPKQK AQLDELSMSE KIAILLIQVG EDTTGEILRH LDIDSITEIS

      51  KQIVQLNGTD KQIGAAVLEE FFAIFQSNQY INTGGLEYAR ELLTRTLGSE

     101  EAKKVMDKLT KSLQTQKNFA YLGKIKPQQL ADFIINEHPQ TIALILAHME

     151  APNAAETLSY FPDEMKAEIS IRMANLGEIS PQVVKRVSTV LENKLESLTS

     201  YKIEVGGLRA VAEIFNRLGQ KSAKTTLARI ESVDNKLAGA IKEMMFTFED

     251  IVKLDNFAIR EILKVADKKD LSLALKTSTK DLTDKFLNNM SSRAAEQFVE

     301  EMQYLGAVKI KDVDVAQRKI IEIVQSLQEK GVIQTGEEED VIE

!!AA_SEQUENCE 1.0
ID   W89910 standard; Protein; 343 AA.
AC   W89910;
DT   18-FEB-1999  (first entry)
DE   Antigen 1 from cluster 35a.
KW   Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW   peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
OS   Helicobacter pylori.
PN   WO9849314-A2.
PD   05-NOV-1998.
PF   27-APR-1998; U08487.
PR   14-OCT-1997; US-061958.
PR   25-APR-1997; US-045107.
PA   (GENE-) GENELABS TECHNOLOGIES INC.
PI   Chow TP, Fry KE, Lim MY, McAtee CP;
DR   WPI; 99-009433/01.
PT   New Helicobacter pylori antigens and related nucleic acid sequences
PT   - useful in serological diagnosis and protective vaccines, providing
PT   long-lasting immune response
PS   Claim 1; Page 243-244; 402pp; English.
CC   The present sequence represents a Helicobacter pylori antigenic protein
CC   that is characterised by immunoreactivity with H. pylori-positive
CC   antisera. The proteins are highly immunogenic and induce a long-lasting
CC   immune response that persists even after antimicriobial treatment. In
CC   antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC   highly sensitive and specific. The specification also describes 69
CC   previously unrecognised immunogenic cluster families. H. pylori antigens
CC   are used to detect H. pylori-specific antibodies, for diagnosing
CC   infection or to confirm eradication of infection, and in vaccines to
CC   protect against H. pylori infection and related diseases (gastritis,
CC   peptic ulcer, gastric adenocarcinoma/lymphoma).
SQ   Sequence   343 AA;

 W89910  Length: 343  February 11, 2000 15:49  Type: P  Check: 8829  ..

       1  MATKLTPKQK AQLNELSMSE KIAILLIQVG EDTTGEILRH LDIDSITEIS

      51  KQIVQLNGTD KQIGAAVLEE FFAIFQSNQY INTGGLEYAR ELLTRTLGSE

     101  EAKKVMDKLT KSLQTQKNFA YLGKIKPPQL ADFIINEHPQ TIALILAHME

     151  APNAAETLSY FPDEMKAEIS IRMANLGEIS PQVLKRVSTV LENKLESLTS

     201  YKIEVGGLRA VAEIFNRLGQ KSAKTTLARI ESVDNKLAGA IKEMMFTFED

     251  IVKLDNFAIR EILKVADKKD LSLALKTSTK DLTDKFLNNM SSRAAEQFVE

     301  EMQYLGAVKI KDVDVAQRKI IEIVQNLQEK GVIQTGEEED VIE

!!AA_SEQUENCE 1.0
ID   W85590 standard; Protein; 354 AA.
AC   W85590;
DT   02-MAR-1999  (first entry)
DE   Polypeptide with 98% homology with human Wnt-11.
KW   Wnt-11; disease; detection; cancer; cardiovascular disease; stroke;
KW   parkinsons disease; schizophrenia; alzheimers disease; acromegaly;
KW   vitelliform macular dystrophy; vitreoretinopathy; assay;
KW   neurological disease.
OS   Homo sapiens.
PN   EP-882791-A2.
PD   09-DEC-1998.
PF   20-MAY-1998; 303992.
PR   12-MAR-1998; GB-005332.
PR   23-MAY-1997; GB-010755.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PI   Barnes MR, Testa TT;
DR   WPI; 99-011648/02.
DR   N-PSDB; V83106.
PT   New human Wnt-11 polypeptide - used to diagnose, treat and prevent
PT   cancer, cardiovascular disease, developmental disorders, stroke and
PT   neurological disease
PS   Claim 1; Page 16-17; 21pp; English.
CC   This polypeptide has at least 98 % sequence identity with the human
CC   Wnt-11 345 amino acid sequence. Host cells are used to produce
CC   this polypeptide from recombinant constructs which is then used
CC   to screen for specific agonists and antagonists (potential therapeutic
CC   agents) and to generate Ab (including induction of a protective immune
CC   response).  Human Wnt-11 polypeptide is involved in developmental
```

CC processes and its agonists and antagonists are used for treatment or
CC prevention of cancer, cardiovascular disease, stroke, neurological
CC disease (e.g. Parkinson's or Alzheimer's diseases, bipolar/unipolar
CC disorders and schizophrenia) or developmental disorders such as
CC acromegaly, vitelliform macular dystrophy or vitreoretinopathy.
CC Fragments of this polynucleotide sequence can be used for isolating
CC related genomic and cDNA clones and for chromosomal or tissue
CC localisation (the Wnt-11 gene is at chromosome 11q13). The specified
CC diseases, or susceptibility to them, are diagnosed by detecting a
CC mutation in a genomic Wnt-11 encoding sequence and/or by analysis for
CC presence or amount of Wnt-11 (or related mRNA) in a patient sample.
SQ Sequence 354 AA;

W85590 Length: 354 February 11, 2000 15:49 Type: P Check: 2306 ..

1 MRARPQVCEA LLFALALQTG VCYGIKWLAL SKTPSALALN QTQHCKQLEG

51 LVSAQVQLCR SNLELMHTVV HAAREVMKAC RRAFADMRWN CSSIELAPNY

101 LLDLERGTRE SAFVYALSAA AISHAIARAC TSGDLPGCSC GPVPGEPPGP

151 GNRWGGCADN LSYGLLMGAK FSDAPMKVKK TGSQANKLMR LHNSEVGRQA

201 LRASLEMKCK CHGVSGSCSI RTCWKGLQEL QDVAADLKTR YLSATKVVHR

251 PMGTRKHLVP KDLDIRPVKD SELVYLQSSP DFCMKNEKVG SHGTQDRQCN

301 KTSNGSDSCD LMCCGRGYNP YTDRVVERCH CKYHWCCYVT CRRCERTVER

351 YVCK

!!AA_SEQUENCE 1.0
ID W85591 standard; Protein; 351 AA.
AC W85591;
DT 02-MAR-1999 (first entry)
DE Polypeptide with 95% homology with human Wnt-11.
KW Wnt-11; disease; detection; cancer; cardiovascular disease; stroke;
KW parkinsons disease; schizophrenia; alzheimers disease; acromegaly;
KW vitelliform macular dystrophy; vitreoretinopathy; assay;
KW neurological disease.
OS Homo sapiens.
PN EP-882791-A2.
PD 09-DEC-1998.
PF 20-MAY-1998; 303992.
PR 12-MAR-1998; GB-005332.
PR 23-MAY-1997; GB-010755.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PI Barnes MR, Testa TT;
DR WPI; 99-011648/02.
DR N-PSDB; V83107.
PT New human Wnt-11 polypeptide - used to diagnose, treat and prevent
PT cancer, cardiovascular disease, developmental disorders, stroke and
PT neurological disease
PS Claim 17; Page 18-19; 21pp; English.
CC This polypeptide has at least 95 % sequence identity with the human
CC Wnt-11 345 amino acid sequence. Host cells are used to produce
CC this polypeptide from recombinant constructs which is then used
CC to screen for specific agonists and antagonists (potential therapeutic
CC agents) and to generate Ab (including induction of a protective immune
CC response). Human Wnt-11 polypeptide is involved in developmental
CC processes and its agonists and antagonists are used for treatment or
CC prevention of cancer, cardiovascular disease, stroke, neurological
CC disease (e.g. Parkinson's or Alzheimer's diseases, bipolar/unipolar
CC disorders and schizophrenia) or developmental disorders such as
CC acromegaly, vitelliform macular dystrophy or vitreoretinopathy.
CC Fragments of this polynucleotide sequence can be used for isolating
CC related genomic and cDNA clones and for chromosomal or tissue
CC localisation (the Wnt-11 gene is at chromosome 11q13). The specified
CC diseases, or susceptibility to them, are diagnosed by detecting a
CC mutation in a genomic Wnt-11 encoding sequence and/or by analysis for
CC presence or amount of Wnt-11 (or related mRNA) in a patient sample.
SQ Sequence 351 AA;

W85591 Length: 351 February 11, 2000 15:49 Type: P Check: 8607 ..

1 RPQVCEAMLF ALALQTGVCY GIKWLALSKT PSALALNQTQ HCKQLEGLVS

51 AQVQLCRSNL ELMHTVVHAA REVMKACRRA FADMRWNCWS IELAPNLLVD

101 LERGTRESAF VYALSAAAIS HAIARACTSG DLPGCSCGPV PGEPPGPGNR

151 WGGCADNLSY WLLMGAKFSD APMKVKKTGS QANKLMGLHN SEVGRQALRA

201 SLEMKCKCHG VSGHCYIGTS WKGLQELQDV AADLKTRYLS ATKVVHRPMG

251 TRKHLVPKDL DIRPVKDSEL VYLQSSPDFC MKNEKVGSHG TQDRQCNKTS

301 NGSDSCDLMC CGRGYNPYTD RVVERCHCKY HWCCYVTCRR CERTVERYVC

351 K

!!AA_SEQUENCE 1.0
ID W89100 standard; Protein; 36 AA.
AC W89100;
DT 01-MAR-1999 (first entry)
DE Polypeptide fragment encoded by gene 206.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9854963-A2.
PD 10-DEC-1998.
PF 04-JUN-1998; U11422.
PR 18-DEC-1997; US-070923.
PR 06-JUN-1997; US-048877.
PR 06-JUN-1997; US-048881.
PR 06-JUN-1997; US-048884.
PR 06-JUN-1997; US-048893.
PR 06-JUN-1997; US-048896.
PR 06-JUN-1997; US-048899.
PR 06-JUN-1997; US-048915.
PR 06-JUN-1997; US-048949.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048972.
PR 06-JUN-1997; US-049020.
PR 06-JUN-1997; US-049375.
PR 05-SEP-1997; US-057628.
PR 05-SEP-1997; US-057635.
PR 05-SEP-1997; US-057644.
PR 05-SEP-1997; US-057647.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057661.
PR 05-SEP-1997; US-057667.
PR 05-SEP-1997; US-057761.
PR 05-SEP-1997; US-057764.
PR 05-SEP-1997; US-057770.
PR 05-SEP-1997; US-057775.
PR 05-SEP-1997; US-057778.
PR 06-JUN-1997; US-048875.
PR 06-JUN-1997; US-048878.
PR 06-JUN-1997; US-048882.
PR 06-JUN-1997; US-048885.
PR 06-JUN-1997; US-048894.
PR 06-JUN-1997; US-048897.
PR 06-JUN-1997; US-048900.
PR 06-JUN-1997; US-048916.
PR 06-JUN-1997; US-048962.
PR 06-JUN-1997; US-048970.
PR 06-JUN-1997; US-048974.
PR 06-JUN-1997; US-049373.
PR 05-SEP-1997; US-057584.

PR 05-SEP-1997; US-057629.
PR 05-SEP-1997; US-057642.
PR 05-SEP-1997; US-057645.
PR 05-SEP-1997; US-057648.
PR 05-SEP-1997; US-057651.
PR 05-SEP-1997; US-057662.
PR 05-SEP-1997; US-057668.
PR 05-SEP-1997; US-057762.
PR 05-SEP-1997; US-057765.
PR 05-SEP-1997; US-057771.
PR 05-SEP-1997; US-057776.
PR 06-JUN-1997; US-048876.
PR 06-JUN-1997; US-048880.
PR 06-JUN-1997; US-048883.
PR 06-JUN-1997; US-048892.
PR 06-JUN-1997; US-048895.
PR 06-JUN-1997; US-048898.
PR 06-JUN-1997; US-048901.
PR 06-JUN-1997; US-048917.
PR 06-JUN-1997; US-048963.
PR 06-JUN-1997; US-048971.
PR 06-JUN-1997; US-049019.
PR 06-JUN-1997; US-049374.
PR 05-SEP-1997; US-057627.
PR 05-SEP-1997; US-057634.
PR 05-SEP-1997; US-057643.
PR 05-SEP-1997; US-057646.
PR 05-SEP-1997; US-057649.
PR 05-SEP-1997; US-057654.
PR 05-SEP-1997; US-057666.
PR 05-SEP-1997; US-057760.
PR 05-SEP-1997; US-057763.
PR 05-SEP-1997; US-057769.
PR 05-SEP-1997; US-057774.
PR 05-SEP-1997; US-057777.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,
PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C,
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW,
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
DR WPI; 99-059865/05.
DR N-PSDB; V84616.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Disclosure; Page 159; 772pp; English.
CC The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a polypeptide fragment encoded by a
CC gene of the invention (see descriptor line for gene number).
SQ Sequence 36 AA;

W89100 Length: 36 February 11, 2000 15:49 Type: P Check: 1342 ..

1 FSVHRPETLF NISRFLLHSL PKDTPSGISK VKILFT

!!AA_SEQUENCE 1.0
ID W88568 standard; Protein; 221 AA.
AC W88568;
DT 01-MAR-1999 (first entry)
DE Secreted protein encoded by gene 35 clone HADAE74.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9854963-A2.
PD 10-DEC-1998.
PF 04-JUN-1998; U11422.
PR 18-DEC-1997; US-070923.
PR 06-JUN-1997; US-048877.
PR 06-JUN-1997; US-048881.
PR 06-JUN-1997; US-048884.
PR 06-JUN-1997; US-048893.
PR 06-JUN-1997; US-048896.
PR 06-JUN-1997; US-048899.
PR 06-JUN-1997; US-048915.
PR 06-JUN-1997; US-048949.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048972.
PR 06-JUN-1997; US-049020.
PR 06-JUN-1997; US-049375.
PR 05-SEP-1997; US-057628.
PR 05-SEP-1997; US-057635.
PR 05-SEP-1997; US-057644.
PR 05-SEP-1997; US-057647.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057661.
PR 05-SEP-1997; US-057667.
PR 05-SEP-1997; US-057761.
PR 05-SEP-1997; US-057764.
PR 05-SEP-1997; US-057770.
PR 05-SEP-1997; US-057775.
PR 05-SEP-1997; US-057778.
PR 06-JUN-1997; US-048875.
PR 06-JUN-1997; US-048878.
PR 06-JUN-1997; US-048882.
PR 06-JUN-1997; US-048885.
PR 06-JUN-1997; US-048894.
PR 06-JUN-1997; US-048897.
PR 06-JUN-1997; US-048900.
PR 06-JUN-1997; US-048916.
PR 06-JUN-1997; US-048962.
PR 06-JUN-1997; US-048970.
PR 06-JUN-1997; US-048974.
PR 06-JUN-1997; US-049373.
PR 05-SEP-1997; US-057584.
PR 05-SEP-1997; US-057629.
PR 05-SEP-1997; US-057642.
PR 05-SEP-1997; US-057645.
PR 05-SEP-1997; US-057648.
PR 05-SEP-1997; US-057651.
PR 05-SEP-1997; US-057662.
PR 05-SEP-1997; US-057668.
PR 05-SEP-1997; US-057762.
PR 05-SEP-1997; US-057765.
PR 05-SEP-1997; US-057771.
PR 05-SEP-1997; US-057776.
PR 06-JUN-1997; US-048876.
PR 06-JUN-1997; US-048880.

PR 06-JUN-1997; US-048883.
PR 06-JUN-1997; US-048892.
PR 06-JUN-1997; US-048895.
PR 06-JUN-1997; US-048898.
PR 06-JUN-1997; US-048901.
PR 06-JUN-1997; US-048917.
PR 06-JUN-1997; US-048963.
PR 06-JUN-1997; US-048971.
PR 06-JUN-1997; US-049019.
PR 06-JUN-1997; US-049374.
PR 05-SEP-1997; US-057627.
PR 05-SEP-1997; US-057634.
PR 05-SEP-1997; US-057643.
PR 05-SEP-1997; US-057646.
PR 05-SEP-1997; US-057649.
PR 05-SEP-1997; US-057654.
PR 05-SEP-1997; US-057666.
PR 05-SEP-1997; US-057760.
PR 05-SEP-1997; US-057763.
PR 05-SEP-1997; US-057769.
PR 05-SEP-1997; US-057774.
PR 05-SEP-1997; US-057777.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,
PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C,
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DW,
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
DR WPI; 99-059865/05.
DR N-PSDB; V84445.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 11; Page 506-507; 772pp; English.
CC The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents human secreted protein (see descriptor
CC line for gene number and clone identification).
SQ Sequence 221 AA;

W88568 Length: 221 February 11, 2000 15:49 Type: P Check: 9823 ..

1 MFHGIPATPG IGAPGNKPEL YEEVKLYKNA REREKYDNMA ELFAVVKTMQ

51 ALEKAYIKDC VSPSEYTAAC SRLLVQYKAA FRQVQGSEIS SIDEFCRKFR

101 LDCPLAMERI KEDRPITIKD DKGNLNRCIA DVVSLFITVM DKLRLEIRAM

151 DEIQPDLREL METMHRMSHL PPDFEGRQTV SQWLQTLSGM SASDELDDSQ

201 VRQMLFDLES AYNAFNRFLH A

!!AA_SEQUENCE 1.0
ID W83391 standard; Protein; 161 AA.
AC W83391;
DT 29-MAR-1999 (first entry)
DE Caenorhabditis elegans synMuv protein LIN-52.
KW LIN-52; synthetic multivulvar; SynMuv; signal transduction;
KW animal model; tumour suppressor; retinoblastoma; cancer; cancer;
KW cell proliferation; gene therapy.
OS Caenorhabditis elegans.
PN WO9854299-A1.
PD 03-DEC-1998.
PF 28-MAY-1998; U11043.
PR 28-MAY-1997; US-047996.
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
PI Ceol C, Horvitz HR, Lu X;
DR WPI; 99-045362/04.
DR N-PSDB; V72864.
PT Novel LIN-37, -35, -55, -52, -53 and -54 gene from C. elegans -
PT useful for treating diseases associated with altered levels of cell
PT proliferation, e.g. carcinomas
PS Claim 7; Fig 11; 70pp; English.
CC This is the amino acid sequence of LIN-52, a novel protein of
CC Caenorhabditis elegans. The lin-52 gene (see V72864) is a novel
CC synthetic multivulvar (synMuv) gene involved in cell fate and cell
CC proliferation, and is part of a pathway that may be used as a
CC genetic and biochemical model system for tumour suppression and
CC cancer in mammals. SynMuv pathway genes and proteins may be used
CC to identify genes which are part of the mammalian pathway and
CC to identify genes, proteins and therapeutic compounds which
CC modulate this pathway. Pure nucleic acids (see V72859-65) encoding
CC C. elegans SynMuv polypeptides selected from LIN-37, -35, -55, -52,
CC -53, -54 and E2F-1 (see W83386-92), are new. Also claimed are:
CC (1) nucleic acids (see V72866-67) encoding the mouse and human
CC homologues of C. elegans LIN-54; (2) vectors containing the nucleic
CC acids; (3) transgenic cells; (4) a pure mammalian SynMuv
CC polypeptide, and (5) an antibody which binds to a SynMuv family
CC protein. The SynMuv nucleic acids and polypeptides can be used to
CC diagnose and treat, especially by gene therapy, conditions
CC involving altered levels of cell proliferation, e.g. SynMuv-
CC associated carcinomas.
SQ Sequence 161 AA;

W83391 Length: 161 February 11, 2000 15:49 Type: P Check: 9212 ..

1 MSRPLGFIGY EFGDDEMFVQ QMIEKKSNAE QAKMLEQQKK MLECTETMPE

51 ESEPVPMKCL DFEEAFQSES VSKGYESPYK NISFLKEDAV TVNTMSHCPA

101 DDIAKLIRNI QNSVYTLGIE EARQCRRGKL LNVLKPTGSA SPRYLQPTPP

151 KNVAEETTGS Q

!!AA_SEQUENCE 1.0
ID W95506 standard; peptide; 446 AA.
AC W95506;
DT 26-MAR-1999 (first entry)
DE Mortierella alpina delta 5 desaturase.
KW Delta 5 desaturase; recombinant; fatty acid desaturase; FAD; PUFA; oil;
KW polyunsaturated fatty acid; linoleic acid; arachidonic acid; linolenic;
KW stearidonic acid; eicosapentaenoic acid; malnutrition; feeding formula;
KW dietary supplement; prostaglandin; restenosis; angioplasty; inflammation;
KW rheumatoid arthritis; psoriasis; osteoporosis; cancer; eczema; AIDS;
KW diabetes; cosmetic; animal feed.
OS Mortierella alpina.
PN WO9846764-A1.
PD 22-OCT-1998.
PF 10-APR-1998; U07421.
PR 24-OCT-1997; US-956985.
PR 11-APR-1997; US-833610.
PR 11-APR-1997; US-834033.
PR 11-APR-1997; US-834655.
PA (ABBO) ABBOTT LAB.
PA (CALJ) CALGENE LLC.

PI Chaudhary S, Huang Y, Knutzon D, Leonard AE, Mukerji P,
PI Thurmond J;
DR WPI; 99-080739/07.
DR N-PSDB; X00891.
PT Nucleic acid construct able to express fatty acid desaturase in
PT plants - useful in human or animal nutrition, as cosmetics and
PT therapeutically, e.g. for restenosis, cancer and diabetes
PS Claim 7; Fig 7A-D; 210pp; English.
CC This represents a Mortierella alpina delta 5 desaturase. The invention
CC relates to a nucleic acid construct that contains at least one of the
CC nucleotide sequences (X00889 to X00891) encoding M. alpina delta 6,
CC delta 12 and delta 5 desaturases (W95504 to W95506) respectively, coupled
CC to an expression control sequence functional in plants. Recombinant plant
CC cells containing at least one DNA encoding a M. alpina fatty acid
CC desaturase (FAD), can be used for the production of polyunsaturated fatty
CC acid (PUFA). These recombinant cells or plants containing them are used
CC to produce oils such as linoleic acid, arachidonic acid, gamma-linolenic
CC acid, dihomo-gamma-linolenic acid, stearidonic acid and eicosapentaenoic
CC acid (EPA). These plant oils are used: (i) to treat malnutrition; (ii) in
CC infant feeding formulas, or dietary supplements or substitutes, for use
CC in humans or animals; (iii) for treating disorders associated with
CC inadequate consumption or production of PUFA (or their metabolites such
CC as prostaglandins), e.g. restenosis after angioplasty, inflammation,
CC AIDS, rheumatoid arthritis, psoriasis, osteoporosis, cancer, eczema,
CC diabetes; (iv) as cosmetics, and (v) as animal feeds. Fragments of the
CC DNA are used as probes to isolate related coding sequences. Recombinant
CC plants can produce high yields of PUFA, since new pathways can be created
CC and unwanted ones suppressed. Plants can be engineered to express oils of
CC particular PUFA composition, e.g. one similar to that in human milk, and
CC product recovery is simpler than with e.g. fish.
SQ Sequence 446 AA;

W95506 Length: 446 February 11, 2000 15:49 Type: P Check: 5663 ..

1 MGTDQGKTFT WEELAAHNTK DDLLLAIRGR VYDVTKFLSR HPGGVDTLLL

51 GAGRDVTPVF EMYHAFGAAD AIMKKYYVGT LVSNELPIFP EPTVFHKTIK

101 TRVEGYFTDR NIDPKNRPEI WGRYALIFGS LIASYYAQLF VPFVVERTWL

151 QVVFAIIMGF ACAQVGLNPL HDASHFSVTH NPTVWKILGA THDFFNGASY

201 LVWMYQHMLG HHPYTNIAGA DPDVSTSEPD VRRIKPNQKW FVNHINQHMF

251 VPFLYGLLAF KVRIQDINIL YFVKTNDAIR VNPISTWHTV MFWGGKAFFV

301 WYRLIVPLQY LPLGKVLLLF TVADMVSSYW LALTFQANHV VEEVQWPLPD

351 ENGIIQKDWA AMQVETTQDY AHDSHLWTSI TGSLNYQAVH HLFPNVSQHH

401 YPDILAIIKN TCSEYKVPYL VKDTFWQAFA SHLEHLRVLG LRPKEE

!!AA_SEQUENCE 1.0
ID W98862 standard; Protein; 304 AA.
AC W98862;
DT 31-MAR-1999 (first entry)
DE H. pylori GHPO 1670 protein.
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease.
OS Helicobacter pylori.
PN WO9843478-A1.
PD 08-OCT-1998.
PF 01-APR-1998; U06371.
PR 29-JUL-1997; US-902615.
PR 01-APR-1997; US-833457.
PR 24-JUN-1997; US-881227.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
DR WPI; 98-542293/46.
DR N-PSDB; X14581.
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
PS Claim 8; Page 1942-1943; 2054pp; English.
CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
SQ Sequence 304 AA;

W98862 Length: 304 February 11, 2000 15:49 Type: P Check: 4789 ..

1 MLESALKYCK EKAIDLLVGF VPKTYSMAQE CNILGLYDDA FIITKQENLV

51 GIISLQGLSY SNLMQKDLEG YFDARQNVLN TISKDIQLRI VAKRRKEFIN

101 QSPNIDNIYA KAIITQFESK GIYKTEYFLV FETITSNVKS FFEKKKLEMT

151 TSINEELEES SKEDKQENEN XSNETHSNTS SKKDKKNKFK KKITFSTKSK

201 RALLIQTIER VKNALKEFKP TLLNSKEVLN FYAEYINGKY IAFNPKLKRL

251 SDTILHLMCI LRKITLSLNF KIKTPFVRVW GLRLMRAKKF LRSLYLLFYT

301 PKLN

!!AA_SEQUENCE 1.0
ID W98755 standard; Protein; 102 AA.
AC W98755;
DT 31-MAR-1999 (first entry)
DE H. pylori GHPO 1075 protein.
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease.
OS Helicobacter pylori.
PN WO9843478-A1.
PD 08-OCT-1998.
PF 01-APR-1998; U06371.
PR 29-JUL-1997; US-902615.
PR 01-APR-1997; US-833457.
PR 24-JUN-1997; US-881227.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
DR WPI; 98-542293/46.
DR N-PSDB; X14474.
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
PS Claim 8; Page 1616; 2054pp; English.
CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
SQ Sequence 102 AA;

W98755 Length: 102 February 11, 2000 15:49 Type: P Check: 3203 ..

1 MIFSFKRTSF CVSVPVLSIK IFFTLARVSK NNASLNTIKG FLNTPITNAM

51 GVARANAQGQ LMTSTLIHTI KAFSKLPPKP NCHNKKLTKA KNNTALEKIS

101 AI

!!AA_SEQUENCE 1.0
ID W98226 standard; Protein; 115 AA.
AC W98226;
DT 31-MAR-1999 (first entry)
DE H. pylori GHPO 1099 protein.
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

KW peptic ulcer disease.
OS Helicobacter pylori.
PN WO9843478-A1.
PD 08-OCT-1998.
PF 01-APR-1998; U06371.
PR 29-JUL-1997; US-902615.
PR 01-APR-1997; US-833457.
PR 24-JUN-1997; US-881227.
PA (HUMA-) HUMAN GENOME SCI INC.
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
DR WPI; 98-542293/46.
DR N-PSDB; X13945.
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
PS Claim 8; Page 159-160; 2054pp; English.
CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
SQ Sequence 115 AA;

W98226 Length: 115 February 11, 2000 15:49 Type: P Check: 2117 ..

1 MKSKITHFIA ISFVLSLFSA CKDEPKKSSQ SHQNNTKITK NNPINQANND

51 IRKIEHEEED EKATKEVNDL INNENKIDEI NNEENADPSQ KRTNNVLQRA

101 TNHQDNLNSP LNRKY

!!AA_SEQUENCE 1.0
ID W88330 standard; Protein; 336 AA.
AC W88330;
DT 26-APR-1999 (first entry)
DE Mannosyl transferase involved in O antigen biosynthesis.
KW O antigen; mannosyl transferase; wbaW gene; diagnosis.
OS Salmonella enterica.
PN WO9850531-A1.
PD 12-NOV-1998.
PF 01-MAY-1998; AU0315.
PR 22-JUL-1997; AU-008162.
PR 01-MAY-1997; AU-006545.
PA (UNSY) UNIV SYDNEY.
PI Reeves PR, Wang L;
DR WPI; 99-059669/05.
DR N-PSDB; X06750.
PT Nucleic acid molecules specific for bacterial polysaccharide
PT antigens - useful for detecting specific strains in, e.g. food,
PT faeces or patient samples
PS Disclosure; Fig 9; 165pp; English.
CC This is the amino acid sequence of a mannosyl transferase that is
CC encoded by the wbaW gene of a gene cluster (see X06750) involved in
CC the biosynthesis of the Salmonella enterica serotype C2 strain M67
CC O antigen. The use of nucleic acid molecules derived from
CC assembly and transport genes, particularly wbd (transferase), wzx
CC (flippase) and wzy (polymerase) genes, within O antigen gene
CC clusters improves the specificity of methods for the detection and
CC identification of O antigens, e.g. in testing food- or faecal-
CC derived samples, or samples from patients. The O antigen is a
CC major virulence factor of S. enterica.
SQ Sequence 336 AA;

W88330 Length: 336 February 11, 2000 15:49 Type: P Check: 7734 ..

1 MNNKKVLMDI SWSNKGGIGR FTDEISKLLC DISKEELYRK CASPLAPLGL

51 AVNIFLRKKT DVVFLPGYIP PLFCSKKFII TIHDLNHLDL NDNSSLFKRL

101 FYNFIIKRGC RKAYKIFTVS NFSKERIVAW SGVNPNKIVT VYNGVSSLFN

151 ADVKPLNLGY KYLLCVGNRK THKNEKCVIS AFAKADIDPS IKLVFTGNPC

201 NDLEKLIIQH GLSERVKFFG FVSEKDLPSL YKGSLGLVFP SLYEGFGLPV

251 VEGMACGIPV LTSLTSSLPE VAGDAAILVD PLSEDAITKG ISRLINDSEL

301 RKHLIQKGLL RAKRFNWQNV VSEIEMVLTE ACDGNK

!!AA_SEQUENCE 1.0
ID Y01440 standard; Protein; 64 AA.
AC Y01440;
DT 18-MAY-1999 (first entry)
DE Secreted protein encoded by gene 58 clone HODAZ55.
KW Human; secreted protein; gene therapy; protein therapy; cancer; weight;
KW tumour; chromosome mapping; forensic; haematological disease; allergy;
KW inflammation; cell proliferation; viral infection; wound healing;
KW modulation; appetite; behaviour; food additive; preservative.
OS Homo sapiens.
PN WO9903990-A1.
PD 28-JAN-1999.
PF 15-JUL-1998; U14613.
PR 18-AUG-1997; US-056361.
PR 16-JUL-1997; US-052661.
PR 16-JUL-1997; US-052870.
PR 16-JUL-1997; US-052871.
PR 16-JUL-1997; US-052872.
PR 16-JUL-1997; US-052873.
PR 16-JUL-1997; US-052874.
PR 16-JUL-1997; US-052875.
PR 22-JUL-1997; US-053440.
PR 22-JUL-1997; US-053441.
PR 22-JUL-1997; US-053442.
PR 18-AUG-1997; US-055683.
PR 18-AUG-1997; US-055724.
PR 18-AUG-1997; US-055725.
PR 18-AUG-1997; US-055726.
PR 18-AUG-1997; US-055946.
PR 18-AUG-1997; US-055952.
PR 18-AUG-1997; US-055985.
PR 18-AUG-1997; US-055989.
PR 18-AUG-1997; US-056359.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Duan R, Feng P, Ferrie AM, Florence KA, Fouad J,
PI Greene JM, Hu J, Ni J, Rosen CA, Ruben SM, Young PE,
PI Yu G;
DR WPI; 99-132234/11.
DR N-PSDB; X22268.
PT New nucleic acids encoding secreted human proteins - potentially
PT useful for treating and diagnosing diseases and identifying specific
PT binding agents
PS Claim 11; Page 232; 251pp; English.
CC The invention relates to nucleic acid sequences (X22211 to X22282)
CC encoding human secreted proteins (Y01383 to Y01454). The secreted protein
CC gene sequences are deposited with the ATCC under deposit number ATCC
CC 209138, 209139 or 209141. Host cells containing vectors comprising the
CC nucleic acid sequences are used for the recombinant expression of the
CC secreted proteins. The polynucleotide and amino acid sequences are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. Pathological conditions can be also diagnosed by
CC determining the amount of the new polypeptides in a sample or by the
CC presence of mutations in the new polynucleotides. The nucleic acid
CC sequences, or its fragments, are useful for chromosome identification and
CC mapping; as antisense and triplex-forming therapeutics; in gene therapy;
CC for (forensic) identification of individuals; as molecular weight
CC markers; to identify related sequences or specific mRNA; in preparation
CC of oligomers and to raise anti-DNA antibodies. Antibodies are useful as
CC immunoassay reagents (including for in vivo imaging) and therapeutically
CC to inhibit or activate particular polypeptides. A very wide range of
CC disorders may be treated with the polynucleotide and polypeptide
CC sequences, e.g. autoimmune or haematological diseases, allergy,
CC inflammation, cancer or other forms of cell proliferation, viral or
CC other infections. The sequences may also be useful in wound healing, to

```
CC   modulate differentiation of embryonic stem cells, to modulate weight,
CC   appetite, behaviour etc. and as food additive or preservative. The
CC   present sequence represents a human secreted protein (see descriptor
CC   line for gene number and clone identification).
SQ   Sequence   64 AA;

 Y01440  Length: 64  February 11, 2000 15:49  Type: P  Check: 9731  ..

       1  MAASVGRATR SAAAHLTQLP PAPRAQRTSP AQPDEGKRRD ADPWRTGPTV

      51  NKTGSIPGRL RGWA

!!AA_SEQUENCE 1.0
ID   W99185 standard; Protein; 314 AA.
AC   W99185;
DT   19-MAY-1999  (first entry)
DE   Rhodococcus corallina ohp operon protein #5.
KW   Rhodococcus corallina; ohp operon; biosensor; mycolic acid bacteria;
KW   inducible promoter; environmental pollutant; industry; medicine.
OS   Rhodococcus corallina.
PN   WO9900517-A2.
PD   07-JAN-1999.
PF   29-JUN-1998; G01893.
PR   27-JUN-1997; GB-013666.
PA   (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
PI   Archer JAC, Powell JAC, Roland HJ, Summers DK;
DR   WPI; 99-095760/08.
DR   N-PSDB; X19362.
PT   Isolating DNA encoding inducible promoter from mycolic acid bacteria
PT   - useful to produce mycolic acid bacterial biosensors for particular
PT   analytes, such as environmental pollutants, e.g. from industry or
PT   medicine
PS   Example 7; Fig 4; 67pp; English.
CC   A method has been developed for identifying and/or isolating DNA from
CC   mycolic acid bacteria which encodes an inducible promoter induced in
CC   response to a particular analyte (and/or associated operon proteins).
CC   The method comprises: (a) culturing source of mycolic acid bacteria in
CC   selective media containing specific analyte and selective for
CC   oligotrophic bacteria; (b) identifying mycolic acid bacteria subsisting
CC   on medium; (c) extracting DNA from these bacteria; (d) incorporating DNA
CC   into vector; (e) cloning vector into suitable host cell; and (f)
CC   screening host cells for inducible promoter and/or proteins to identify
CC   vectors encoding it. The method allows isolation of DNA encoding a
CC   promoter which is induced in response to a particular analyte (and/or
CC   associated operon proteins), which can be used to produce biosensors for
CC   the analyte. The DNA is incorporated into vectors which are used to
CC   transform host cells to produce the biosensors. It is especially useful
CC   when the analyte is an environmental pollutant (e.g. from industry or
CC   medicine), especially a hydrophic organic compound such as components of
CC   fuels, pesticides; the biosensors may then be used to detect small
CC   concentrations of analytes in samples e.g. for pollution monitoring. The
CC   identified DNA can also be modified to produce modified inducible
CC   promoters and/or operons also useful in biosensor production. The method
CC   allows rapid isolation of promoters (and/or operon proteins) whist
CC   minimising host restrictions (e.g. thick cell walls of mycolic acid
CC   bacteria which confer resistance to cell lysis) and requiring no
CC   knowledge of inducible enzyme chemistry involved (which may be unknown
CC   for specific analyte). The present sequence represents a protein from
CC   the R. corallina ohp operon, given in the present invention.
SQ   Sequence   314 AA;

 W99185  Length: 314  February 11, 2000 15:49  Type: P  Check: 9793  ..

       1  MPVALCAMSH SPLMGRNDPE QEVIDAVDAA FDHARRFVAD FAPDLIVIFA

      51  PDHYNGVFYD LLPPFCIGAA AQSVGDYGTE AGPLDVDRDA AYAVARDVLD

     101  SGIDVAFSER MHVDHGFAQA LQLLVGSITA VPTVPIFINS VAEPLGPVSR

     151  VRLLGEAVGR AAAKLDKRVL FVGSGGLSHD PPVPQFATAP EEVRERLIDG

     201  RNPSAAERDA REQRVITAGR DFAAGTAAIQ PLNPEWDRHL LDVLASGDLE

     251  QIDAWTNDWF VEQAGHSSHE VRTWIAAYAA MSAAGKYRVT STFYREIHEW

     301  IAGFGITTAV AVDE

!!AA_SEQUENCE 1.0
ID   W97722 standard; Protein; 302 AA.
AC   W97722;
DT   21-MAY-1999  (first entry)
DE   Staphylococcus aureus mutant P15B9 virulence gene product.
KW   Virulence; P15B9; vaccine; antibacterial; antibiotic; screening;
KW   bacteraemia.
OS   Staphylococcus aureus.
PN   WO9901473-A2.
PD   14-JAN-1999.
PF   03-JUL-1998; G01974.
PR   03-JUL-1997; US-887534.
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
PI   Holden DW;
DR   WPI; 99-105999/09.
DR   N-PSDB; X07134.
PT   Inhibition of virulence genes from Staphylococcus aureus - useful
PT   for, e.g. screening for potential anti-microbial agents
PS   Claim 1; Page 179-180; 203pp; English.
CC   This is the deduced amino acid sequence of the protein product of a
CC   virulence gene (see X07134) newly identified in Staphylococcus aureus
CC   mutant P15B9.  S. aureus mutants containing chromosomal insertions
CC   of a signature-tagged transposon were generated, and those with
CC   attenuated virulence were identified in a mouse model of bacteraemia.
CC   The nucleotide sequences of regions flanking the transposon insertion
CC   sites of these mutants were determined, and database sequence
CC   comparisons were performed to identify these virulence genes (see
CC   X07088-136) and the possible function of their protein products
CC   (see W97680-724).  The P15B9 virulence gene product is unknown but
CC   shows 41% identity to the hypothetical 33.7 kDa protein YHCT of
CC   Bacillus subtilis p54604.  A claimed method of identifying an
CC   antibacterial agent involves assaying potential agents for the
CC   ability to interfere with the expression of S. aureus virulence
CC   gene products.  Also new is a S. aureus organism containing a
CC   functional mutation in one of the virulence genes, and its use in
CC   vaccine compositions.
SQ   Sequence   302 AA;

 W97722  Length: 302  February 11, 2000 15:49  Type: P  Check: 1929  ..

       1  MNQKGRGLEI LINEKQDGQW LFSVLKTALK ASKPVIQDWM SHQQIKVNHE

      51  SVLNNMIVKK GDRVFIDLQE SEASSVIPEY GELDILFEDN HMLIINKPAG

     101  IATHPNEDGQ TGTLANLIAY HYQINGETCK VRHVHRLDQD TSGAIVFAKH

     151  RLAHAILDQQ LEKKTLKRTY TAIAEGKLRT KKGTINPPIG RDRSHPTRRR

     201  VSPGGQTAVT HFKVMASNAK ERLSLVELEL ETGRTHQIRV HLASLGHPLT

     251  GDSLYGGGSK LLNRQALHAN KVQAVHPITD ELIVAEAPFP ADMKNLCRTY

     301  FS

!!AA_SEQUENCE 1.0
ID   Y11009 standard; Protein; 142 AA.
AC   Y11009;
DT   08-JUN-1999  (first entry)
DE   H. pylori ORF hp6p22217_23564012_f1_5 secreted protein.
KW   Vaccine; probe; diagnostic; ORF; cell envelope protein;
KW   secreted protein; cellular protein.
OS   Helicobacter pylori.
PN   WO9818323-A1.
PD   07-MAY-1998.
PF   28-OCT-1997; U19575.
PR   14-JUL-1997; US-891928.
PR   28-OCT-1996; US-739150.
PR   06-DEC-1996; US-759739.
PA   (ASTR ) ASTRA AB.
```

PI Alm RA, Smith D;
DR WPI; 98-271811/24.
DR N-PSDB; X30476.
PT Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
PT infections
PS Claims 27, 31; Page 219; 279pp; English.
CC Recombinant or substantially pure preparations of H. pylori polypeptides
CC are disclosed, together with the nucleic acids encoding them. In all,
CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
CC secreted proteins or other cellular proteins. Vaccines containing the
CC nucleic acids or proteins are claimed, as are probes containing at least
CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful
CC for treating or reducing the risk of H. pylori infections, and the
CC probes can be used diagnostically for detecting the presence of
CC Helicobacter in a sample. The products are also of use in screening
CC for compounds having the ability to interfere with the H. pylori life
CC cycle or to inhibit H. pylori infection.
SQ Sequence 142 AA;

Y11009 Length: 142 February 11, 2000 15:49 Type: P Check: 2840 ..

1 MKTNFYKIKL LFAWCLIIGM FNAPLNADQN TDIKDISPED MALNSVGLVS

51 RDQLKIEIPK ETLEQKVAIL NDYNDKNVNI KFDDISLGSF QPNDNLGINA

101 MWGIQNLLMS QMMSNYGPNN SFMYGYAPTY SDSSFLPPIL GY

!!AA_SEQUENCE 1.0
ID Y11062 standard; Protein; 116 AA.
AC Y11062;
DT 08-JUN-1999 (first entry)
DE H. pylori ORF 06cp30603_10744075_c3_136 secreted protein.
KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
KW secreted protein; cytoplasmic protein; cellular protein.
OS Helicobacter pylori.
PN WO9824475-A1.
PD 11-JUN-1998.
PF 05-DEC-1997; U22104.
PR 14-JUL-1997; US-891928.
PR 05-DEC-1996; US-759625.
PR 25-MAR-1997; US-823745.
PA (ASTR) ASTRA AB.
PI Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;
DR WPI; 98-333051/29.
DR N-PSDB; X30591.
PT New isolated Helicobacter pylori nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of infection by
PT H. pylori and other Helicobacter species
PS Claims 37, 41; Page 224-225; 339pp; English.
CC Recombinant or substantially pure preparations of H. pylori polypeptides
CC are disclosed, together with the nucleic acids encoding them. In all,
CC 97 ORFs are shown. The proteins are variously cell envelope proteins,
CC cytoplasmic proteins, secreted proteins or other cellular proteins.
CC Vaccines containing the nucleic acids or proteins are claimed, as are
CC probes containing at least 8 nucleotides from the nucleic acid
CC sequences. The vaccines are useful for treating or reducing the risk of
CC H. pylori infections, and the probes can be used diagnostically for
CC detecting the presence of Helicobacter in a sample. The products are
CC also of use in screening for compounds having the ability to interfere
CC with the H. pylori life cycle or to inhibit H. pylori infection.
SQ Sequence 116 AA;

Y11062 Length: 116 February 11, 2000 15:49 Type: P Check: 4555 ..

1 LMKSKITHFI VISFVLSVLS ACKDEPKKSS QSHQNNTKTT QNNQINQPNK

51 DIKKIEHEEE DEKVTKEVND LINNENKIDE INNEENADPS QKRTNNVLQR

101 ATNHQDNLSS PLNRKY

!!AA_SEQUENCE 1.0
ID W97887 standard; Protein; 375 AA.
AC W97887;
DT 07-JUN-1999 (first entry)
DE Human myostatin.
KW Myostatin; human; transforming growth factor beta;
KW double muscling; muscle hyperplasia; transgenic animal.
OS Homo sapiens.
PN WO9902667-A1.
PD 21-JAN-1999.
PF 14-JUL-1998; IB1197.
PR 15-JAN-1998; US-007761.
PR 14-JUL-1997; US-891789.
PA (UYLI-) UNIV LIEGE.
PI Georges M, Grobet L, Poncelet D;
DR WPI; 99-120869/10.
PT Increasing muscle mass in mammals - by decreasing myostatin
PT expression
PS Disclosure; Page 65; 75pp; English.
CC This is the amino acid sequence of human myostatin, a member of
CC the transforming growth factor-beta superfamily. The invention
CC relates to factors affecting muscle development in mammals,
CC especially to cloning of the myostatin gene and determining
CC the role of the gene in muscle hyperplasia. A mutation of the
CC bovine myostatin gene (see X24415-16) has been detected. Certain
CC breeds of cattle homozygous for the mutant gene are double-muscled.
CC A new method of increasing muscle mass of a mammal having myostatin-
CC expressing muscle cells, comprises administration of a nucleic acid
CC molecule substantially complementary to at least a portion of mRNA
CC encoding myostatin and of sufficient length to reduce myostatin
CC expression and thus increase muscle mass. A ribozyme may also be
CC used. Also claimed are: a method for determining muscular
CC hyperplasia (MH) in a mammal using primers based upstream and
CC downstream of the mutation; a diagnostic kit for determining the
CC genotype of a sample of genetic material; a method for determining
CC MH in a mammal; a method for determining double muscling in a
CC bovine; a method for determining the myostatin genotype of an
CC animal; purified myostatin; isolated nucleic acids; a microbial
CC host cell; a probe based on the myostatin gene mutation; transgenic
CC mammals having MH phenotype; and a myostatin knockout animal.
CC Primers are preferably based on genomic bovine myostatin DNA (see
CC X24464) and human myostatin cDNA (see X24418).
SQ Sequence 375 AA;

W97887 Length: 375 February 11, 2000 15:49 Type: P Check: 1814 ..

1 MQKLQLCVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACTWRQNTKS

51 SRIEAIKIQI LSKLRLETAP NISKDVIRQL LPKAPPLREL IDQYDVQRDD

101 SSDGSLEDDD YHATTETIIT MPTESDFLMQ VDGKPKCCFF KFSSKIQYNK

151 VVKAQLWIYL RPVETPTTVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

201 GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPGPGEDGLN

251 PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

301 PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRGSAGPCC TPTKMSPINM

351 LYFNGKEQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID W97884 standard; Protein; 375 AA.
AC W97884;
DT 07-JUN-1999 (first entry)
DE Bovine myostatin.
KW Myostatin; cattle; bovine; transforming growth factor beta;
KW double muscling; muscle hyperplasia; transgenic animal.
OS Bos taurus.
PN WO9902667-A1.
PD 21-JAN-1999.
PF 14-JUL-1998; IB1197.
PR 15-JAN-1998; US-007761.
PR 14-JUL-1997; US-891789.

PA (UYLI-) UNIV LIEGE.
PI Georges M, Grobet L, Poncelet D;
DR WPI; 99-120869/10.
DR N-PSDB; X24415, X24464.
PT Increasing muscle mass in mammals - by decreasing myostatin
PT expression
PS Claim 19; Page 55; 75pp; English.
CC This is the amino acid sequence of bovine myostatin, a member of
CC the transforming growth factor beta superfamily, as encoded by the
CC wild-type gene (see X24415). A mutation of this gene (see X24416)
CC has been detected in cattle. Cattle of the Belgian Blue breed
CC homozygous for the mutant gene are double-muscled. A new method of
CC increasing muscle mass of a mammal having myostatin-expressing
CC muscle cells, comprises administration of a nucleic acid molecule
CC substantially complementary to at least a portion of mRNA
CC encoding myostatin and of sufficient length to reduce myostatin
CC expression and thus increase muscle mass. A ribozyme may also be
CC used. Also claimed are: a method for determining muscular
CC hyperplasia (MH) in a mammal using primers based upstream and
CC downstream of the mutation; a diagnostic kit for determining
CC the genotype of a sample of genetic material; a method for
CC determining MH in a mammal; a method for determining double
CC muscling in a bovine animal; a method for determining the myostatin
CC genotype of an animal; purified myostatin; isolated nucleic acids;
CC a microbial host cell; a probe based on the myostatin gene
CC mutation; transgenic mammals having MH phenotype; a myostatin
CC knockout animal; and a transgenic bovine having a gene encoding
CC active myostatin, and a heterologous nucleotide sequence antisense
CC to that gene, and optionally further containing a gene encoding a
CC nucleic acid sequence with ribozyme activity in transcriptional
CC association with the antisense sequence.
SQ Sequence 375 AA;

W97884 Length: 375 February 11, 2000 15:49 Type: P Check: 8371 ..

1 MQKLQISVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACLWRENTTS

51 SRLEAIKIQI LSKLRLETAP NISKDAIRQL LPKAPPLLEL IDQFDVQRDA

101 SSDGSLEDDD YHARTETVIT MPTESDLLTQ VEGKPKCCFF KFSSKIQYNK

151 LVKAQLWIYL RPVKTPATVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

201 GICQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPEPGEDGLT

251 PFLEVKVTDT PKRSRRDFGL DCDEHSTESR CCRYPLTVDF EAFGWDWIIA

301 PKRYKANYCS GECEFVFLQK YPHTHLVHQA NPRASAGPCC TPTKMSPINM

351 LYFNGEGQII YGKIPAMVVD RCGCS

!!AA_SEQUENCE 1.0
ID W97885 standard; Protein; 286 AA.
AC W97885;
DT 07-JUN-1999 (first entry)
DE Bovine myostatin (mutant form).
KW Myostatin; cattle; bovine; transforming growth factor beta;
KW double muscling; muscle hyperplasia; transgenic animal; mutant.
OS Bos taurus.
PN WO9902667-A1.
PD 21-JAN-1999.
PF 14-JUL-1998; IB1197.
PR 15-JAN-1998; US-007761.
PR 14-JUL-1997; US-891789.
PA (UYLI-) UNIV LIEGE.
PI Georges M, Grobet L, Poncelet D;
DR WPI; 99-120869/10.
DR N-PSDB; X24416.
PT Increasing muscle mass in mammals - by decreasing myostatin
PT expression
PS Disclosure; Page 57; 75pp; English.
CC This is the amino acid sequence of a mutant bovine myostatin,
CC which is truncated when compared to the wild-type polypeptide (see
CC W97884). The mutation results from an 11-bp deletion, i.e.
CC nt821del(11), in the coding region (see X24416) of the myostatin
CC gene. Cattle of the Belgian Blue breed homozygous for the mutant
CC gene are double-muscled. A new method of increasing muscle mass of
CC a mammal having myostatin-expressing muscle cells, comprises
CC administration of a nucleic acid molecule substantially
CC complementary to at least a portion of mRNA encoding myostatin and
CC of sufficient length to reduce myostatin expression and thus
CC increase muscle mass. Also claimed are: a method for determining
CC muscular hyperplasia (MH) in a mammal using primers based upstream
CC and downstream of the mutation; a diagnostic kit for determining
CC the genotype of a sample of genetic material; a method for
CC determining MH in a mammal; a method for determining double
CC muscling in a bovine animal; a method for determining the myostatin
CC genotype of an animal; purified myostatin; isolated nucleic acids;
CC a microbial host cell; a probe based on the myostatin gene
CC mutation; transgenic mammals having MH phenotype; a myostatin
CC knockout animal; and a transgenic bovine having a gene encoding
CC active myostatin, and a heterologous nucleotide sequence antisense
CC to that gene, and optionally further containing a gene encoding a
CC nucleic acid sequence with ribozyme activity in transcriptional
CC association with the antisense sequence.
SQ Sequence 286 AA;

W97885 Length: 286 February 11, 2000 15:49 Type: P Check: 1462 ..

1 MQKLQISVYI YLFMLIVAGP VDLNENSEQK ENVEKEGLCN ACLWRENTTS

51 SRLEAIKIQI LSKLRLETAP NISKDAIRQL LPKAPPLLEL IDQFDVQRDA

101 SSDGSLEDDD YHARTETVIT MPTESDLLTQ VEGKPKCCFF KFSSKIQYNK

151 LVKAQLWIYL RPVKTPATVF VQILRLIKPM KDGTRYTGIR SLKLDMNPGT

201 GIWQSIDVKT VLQNWLKQPE SNLGIEIKAL DENGHDLAVT FPEPGEDGLT

251 PFLEVKVTDT PKRSRRDFGL DCDRISMLSL PSNCGF

!!AA_SEQUENCE 1.0
ID W97886 standard; Protein; 376 AA.
AC W97886;
DT 07-JUN-1999 (first entry)
DE Murine myostatin.
KW Myostatin; mouse; transforming growth factor beta;
KW double muscling; muscle hyperplasia; transgenic animal.
OS Mus sp.
PN WO9902667-A1.
PD 21-JAN-1999.
PF 14-JUL-1998; IB1197.
PR 15-JAN-1998; US-007761.
PR 14-JUL-1997; US-891789.
PA (UYLI-) UNIV LIEGE.
PI Georges M, Grobet L, Poncelet D;
DR WPI; 99-120869/10.
DR N-PSDB; X24417.
PT Increasing muscle mass in mammals - by decreasing myostatin
PT expression
PS Disclosure; Page 60; 75pp; English.
CC This is the amino acid sequence of murine myostatin, a member of
CC the transforming growth factor beta superfamily. The invention
CC relates to factors affecting muscle development in mammals,
CC including the detection of a mutation in the bovine myostatin
CC gene (see X24415-16). Cattle of the Belgian Blue breed homozygous
CC for the mutant gene are double-muscled. A new method of increasing
CC muscle mass of a mammal having myostatin-expressing muscle cells,
CC comprises administration of a nucleic acid molecule substantially
CC complementary to at least a portion of mRNA encoding myostatin
CC (including murine myostatin) and of sufficient length to reduce
CC myostatin expression and thus increase muscle mass. A ribozyme may
CC also be used. Also claimed are: a method for determining muscular
CC hyperplasia (MH) in a mammal using primers based upstream and
CC downstream of the mutation; a diagnostic kit for determining
CC the genotype of a sample of genetic material; a method for

CC determining MH in a mammal; a method for determining double
CC muscling in a bovine animal; a method for determining the myostatin
CC genotype of an animal; purified myostatin; isolated nucleic acids;
CC a microbial host cell; a probe based on the myostatin gene
CC mutation; transgenic mammals having MH phenotype; and a myostatin
CC knockout animal; and a transgenic bovine having a gene encoding
CC active myostatin.
SQ Sequence 376 AA;

W97886 Length: 376 February 11, 2000 15:49 Type: P Check: 2420 ..

1 MMQKLQMYVY IYLFMLIAAG PVDLNEGSER EENVEKEGLC NACAWRQNTR

51 YSRIEAIKIQ ILSKLRLETA PNISKDAIRQ LLPRAPPLRE LIDQYDVHRD

101 DSSDGSLEDD DYHATTETII TMPTESDFLM QADGKPKCCF FKFSSKIQYN

151 KVVKAQLWIY LRPVKTPTTV FVQILRLIKP MKDGTRYTGI RSLKLDMSPG

201 TGIWQSIDVK TVLQNWLKQP ESNLGIEIKA LDENGHDLAV TFPGPGEDGL

251 NPFLEVKVTD TPKRSRRDFG LDCDEHSTES RCCRYPLTVD FEAFGWDWII

301 APKRYKANYC SGECEFVFLQ KYPHTHLVHQ ANPRGSAGPC CTPTKMSPIN

351 MLYFNGKEQI IYGKIPAMVV DRSGCS

!!AA_SEQUENCE 1.0
ID W97840 standard; Protein; 207 AA.
AC W97840;
DT 07-JUN-1999 (first entry)
DE Winged bean chymotrypsin inhibitor WCI-3.
KW Chymotrypsin inhibitor; WCI-3; trypsin inhibitor; WTI-1B;
KW protease inhibitor; enzyme engineering; protein engineering;
KW artificially synthesised gene; winged bean; rice; transgenic plant;
KW insect resistance.
OS Psophocarpus tetragonolobus.
FH Key Location/Qualifiers
FT Peptide 1. .24
FT /note= "hydrophilic peptide utilised in
FT synthetic trypsin inhibitor"
PN EP-900842-A2.
PD 10-MAR-1999.
PF 01-SEP-1998; 307004.
PR 01-SEP-1997; JP-236332.
PA (NORQ) TOHOKU NAT AGRIC EXPERIMENT STATION.
PI Mochizuki A;
DR WPI; 99-155935/14.
DR N-PSDB; X24283.
PT Artificial synthetic trypsin inhibitor gene (WTI-1b) useful for
PT insect resistance in transgenic plants - designed on winged bean
PT chemotrypsin inhibitor (WCI-3) using codon and base comparisons
PS Disclosure; Page 15; 21pp; English.
CC This polypeptide comprises chymotrypsin inhibitor WCI-3 of winged
CC bean (Psophocarpus tetragonolobus (L.) DC.). An artificially
CC synthesised gene (see X24282) is claimed which codes for a trypsin
CC inhibitor (see W97839) designed for stable expression in rice.
CC The trypsin inbihitor comprises the winged bean WTI-1B polypeptide
CC having an N-terminal hydrophilic peptide derived from WCI-3 (i.e.
CC amino acid residues 1-24 of WCI-3) utilised for transporting the
CC protein precursor to a vacuole after translation. The synthetic
CC was obtained by: (i) comparing WTI-1B and WCI-3 amino acid
CC sequences, and obtaining homologous amino acid pairs (one each
CC from WTI-1a and b); and (ii) selecting codons from WCI-3 if the
CC amino acid pair is the same, and from WTI-1B if they are
CC different. High usage codons are selected and sequences which
CC cause mRNA stability are modified (see also X24294-96). Also
CC claimed is a plant containing the artificially synthesised gene.
CC Trangenic plants can be produced that are resistant to pests,
CC especially lepidopteran insects.
SQ Sequence 207 AA;

W97840 Length: 207 February 11, 2000 15:49 Type: P Check: 7697 ..

1 MKSTTFLALF LLSAIISHLP SSTADDDLVD AEGNLVENGG TYYLLPHIWA

51 HGGGIETAKT GNEPCPLTVV RSPNEVSKGE PIRISSQFLS LFIPRGSLVA

101 LGFANPPSCA ASPWWTVVDS PQGPAVKLSQ QKLPEKDILV FKFEKVSHSN

151 IHVYKLLYCQ HDEEDVKCDQ YIGIHRDRNG NRRLVVTEEN PLELVLLKAK

201 SETASSH

!!AA_SEQUENCE 1.0
ID Y03794 standard; Protein; 144 AA.
AC Y03794;
DT 11-JUN-1999 (first entry)
DE S. aureus polypeptide.
KW Staphylococcus aureus polypeptide; thyroiditis; infective carditis;
KW lung abscess; secretory diarrhoea; cerebral abscess; conjunctivitis;
KW toxic shock syndrome; folliculitis; septic arthritis; antibacterial;
KW H pylori infection; gastric ulcer; adenocarcinoma.
OS Staphylococcus aureus.
PN EP-905243-A2.
PD 31-MAR-1999.
PF 03-AUG-1998; 306185.
PR 05-AUG-1997; US-055387.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PI Burnham MKR, Lonetto MA, Warren PV;
DR WPI; 99-192667/17.
DR N-PSDB; X31864.
PT New essential polypeptides from Staphylococcus aureus useful for
PT treating diseases such as infective endocarditis and toxic shock
PT syndrome
PS Claim 31; Page 61-62; 70pp; English.
CC The invention provides new Staphylococcus aureus polypeptides (Y03781-94)
CC and the genes (X31851-864) encoding them. Host cells containing vectors
CC comprising the nucleic acid sequences are used for the recombinant
CC expression of the proteins. The polypeptides can be used to screen for
CC modulators for use in antibacterial therapy. The polypeptides, their
CC antagonists and agonists are used to prevent or treat diseases caused by
CC S. aureus such as thyroiditis, lung abscesses, infective carditis,
CC secretory diarrhoea, cerebral abscesses, conjunctivitis, toxic shock
CC syndrome folliculitis and septic arthritis. Screening for the presence of
CC the polypeptides may be used to diagnose, predict the susceptibility to,
CC or stage the progress of these S. aureus diseases and diseases caused by
CC Helicobacter pylori such as gastric ulcers and gastric adenocarcinoma.
CC There is not much information known about the essential genes expressed
CC by S. aureus during infection but these new polypeptides have been
CC identified as essential. They can therefore be used to develop
CC antibacterial compounds specific for those essential genes and this
CC ensures the effectiveness of the compounds in killing S. aureus. In
CC addition, these polypeptides can be used to effectively diagnose and
CC treat infections and diseases caused by S. aureus without the risk of
CC development of antibiotic resistance. The present sequence represents a
CC S. aureus polypeptide which has homology to an apparent transcriptional
CC regulator.
SQ Sequence 144 AA;

Y03794 Length: 144 February 11, 2000 15:49 Type: P Check: 7277 ..

1 MDRTKQSLNV FVGMNRALDT LEQITKEDVK RYGLNITEFA VLELLYNKGP

51 QPIQRIRDRV LIASSSISYV VSQLEDKGWI TREKDKDDKR VYMACLTEKG

101 QSQMADIFPK HAETLTKAFD VLTKDELTIL QQAFKKLSAQ STEV

!!AA_SEQUENCE 1.0
ID Y03782 standard; Protein; 135 AA.
AC Y03782;
DT 11-JUN-1999 (first entry)
DE S. aureus polypeptide.
KW Staphylococcus aureus polypeptide; thyroiditis; infective carditis;
KW lung abscess; secretory diarrhoea; cerebral abscess; conjunctivitis;

KW toxic shock syndrome; folliculitis; septic arthritis; antibacterial;
KW H pylori infection; gastric ulcer; adenocarcinoma.
OS Staphylococcus aureus.
PN EP-905243-A2.
PD 31-MAR-1999.
PF 03-AUG-1998; 306185.
PR 05-AUG-1997; US-055387.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PI Burnham MKR, Lonetto MA, Warren PV;
DR WPI; 99-192667/17.
DR N-PSDB; X31852.
PT New essential polypeptides from Staphylococcus aureus useful for
PT treating diseases such as infective endocarditis and toxic shock
PT syndrome
PS Claim 11; Page 24-25; 70pp; English.
CC The invention provides new Staphylococcus aureus polypeptides (Y03781-94)
CC and the genes (X31851-864) encoding them. Host cells containing vectors
CC comprising the nucleic acid sequences are used for the recombinant
CC expression of the proteins. The polypeptides can be used to screen for
CC modulators for use in antibacterial therapy. The polypeptides, their
CC antagonists and agonists are used to prevent or treat diseases caused by
CC S. aureus such as thyroiditis, lung abscesses, infective carditis,
CC secretory diarrhoea, cerebral abscesses, conjunctivitis, toxic shock
CC syndrome folliculitis and septic arthritis. Screening for the presence of
CC the polypeptides may be used to diagnose, predict the susceptibility to,
CC or stage the progress of these S. aureus diseases and diseases caused by
CC Helicobacter pylori such as gastric ulcers and gastric adenocarcinoma.
CC There is not much information known about the essential genes expressed
CC by S. aureus during infection but these new polypeptides have been
CC identified as essential. They can therefore be used to develop
CC antibacterial compounds specific for those essential genes and this
CC ensures the effectiveness of the compounds in killing S. aureus. In
CC addition, these polypeptides can be used to effectively diagnose and
CC treat infections and diseases caused by S. aureus without the risk of
CC development of antibiotic resistance. The present sequence represents a
CC S. aureus polypeptide which has homology to a B. subtilis probable
CC acetyltransferase.
SQ Sequence 135 AA;

Y03782 Length: 135 February 11, 2000 15:49 Type: P Check: 881 ..

1 IRHERAFIKR GLLQFADGNG FQCGIWYEGT LVGVIGLHEI NHMHRKTSLG

51 YYLYKQFEGH GIMTQALEAL INYCFYKPYF ISIYITASFH FLINSIYFLT

101 RLDLFPHSIF IFPSHLSTWL FISSTLHLSL LIFSF

!!AA_SEQUENCE 1.0
ID Y02934 standard; Protein; 53 AA.
AC Y02934;
DT 11-JUN-1999 (first entry)
DE Fragment of human secreted protein encoded by gene 106.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
PN WO9902546-A1.
PD 21-JAN-1999.
PF 07-JUL-1998; U13684.
PR 12-SEP-1997; US-058785.
PR 08-JUL-1997; US-051916.
PR 08-JUL-1997; US-051918.
PR 08-JUL-1997; US-051919.
PR 08-JUL-1997; US-051920.
PR 08-JUL-1997; US-051925.
PR 08-JUL-1997; US-051926.
PR 08-JUL-1997; US-051928.
PR 08-JUL-1997; US-051929.
PR 08-JUL-1997; US-051930.
PR 08-JUL-1997; US-051931.
PR 08-JUL-1997; US-051932.
PR 08-JUL-1997; US-052732.
PR 08-JUL-1997; US-052733.
PR 08-JUL-1997; US-052793.
PR 08-JUL-1997; US-052795.
PR 08-JUL-1997; US-052803.
PR 18-AUG-1997; US-055684.
PR 18-AUG-1997; US-055722.
PR 18-AUG-1997; US-055723.
PR 18-AUG-1997; US-055947.
PR 18-AUG-1997; US-055948.
PR 18-AUG-1997; US-055949.
PR 18-AUG-1997; US-055950.
PR 18-AUG-1997; US-055953.
PR 18-AUG-1997; US-055954.
PR 18-AUG-1997; US-055964.
PR 18-AUG-1997; US-055984.
PR 18-AUG-1997; US-056360.
PR 12-SEP-1997; US-058660.
PR 12-SEP-1997; US-058661.
PR 12-SEP-1997; US-058664.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Brewer LA, Ebner R, Fischer CL, Kyaw H, Lafleur DW, Li Y, Moore PA,
PI Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR, Zeng Z;
DR WPI; 99-120770/10.
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Disclosure; Page 122; 464pp; English.
CC This sequence represents a fragment of a secreted human protein encoded
CC by the nucleic acid molecule detailed in the descriptor line. The gene
CC can be used to generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. X27302) for increasing the stability of
CC the fused protein as compared to the human protein only.
CC The invention relates to 123 novel genes and their fragments (nucleic
CC acid sequences: X27311-X27449; amino acid sequences Y02650-Y02788) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 123 polynucleotides, based on
CC which tissues they are most highly expressed in (see X27311 for described
CC uses).
SQ Sequence 53 AA;

Y02934 Length: 53 February 11, 2000 15:49 Type: P Check: 2639 ..

1 TLVAGSPCSL SRWIMAGFCH GELVQSDMES QEWERGQVVL SHTSLPWCYV

51 SPR

!!AA_SEQUENCE 1.0
ID Y11970 standard; Protein; 113 AA.
AC Y11970;
DT 18-JUN-1999 (first entry)
DE Human 5' EST secreted protein SEQ ID No: 570.
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
OS Homo sapiens.
PN WO9906550-A2.
PD 11-FEB-1999.
PF 31-JUL-1998; IB1232.
PR 01-AUG-1997; US-905144.
PA (GEST) GENSET.
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
DR WPI; 99-153780/13.

DR N-PSDB; X40692.
PT New isolated prostate-derived nucleic acids - used to develop
PT products which may have cytokine, immune regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity
PS Claim 34; Page 660; 675pp; English.
CC X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins expressed in prostate, and encode the proteins given in
CC Y11716 to Y11993 respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell proliferation and
CC differentiation activity, haematopoiesis regulating activity, tissue
CC growth regulating activity, reproductive hormone regulating activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC activity or other activities. The products can be used in forensic, gene
CC therapy and chromosome mapping procedures. The sequences can also be used
CC for obtaining corresponding promoter sequences. The nucleic acids
CC encoding the signal peptides can be used for directing extracellular
CC secretion of a polypeptide or the insertion of a polypeptide into a
CC membrane, or importing a polypeptide into a cell.
SQ Sequence 113 AA;

Y11970 Length: 113 February 11, 2000 15:49 Type: P Check: 1817 ..

1 MANLFIRKMV NPLLYLSRHT VKPRALSTXL FGSIRGAAPV AVEPGAAVRS

51 LLSPGLLPHL LPALGFKNKT VLKKRCKDCY LVKRRGRWYV YCKTHPRHKQ

101 RHMXTLSLQS HAQ

!!AA_SEQUENCE 1.0
ID W96262 standard; Protein; 423 AA.
AC W96262;
DT 14-JUN-1999 (first entry)
DE Brn-3a polypeptide.
KW Brn-3a; Bcl-2; neurons; neuronal cells; apoptosis; cell death; CNS;
KW PNS; central nervous system; parasympathetic nervous system;
KW development; injury; neurotrophic factor; nerve growth factor; NGF;
KW ciliary neurotrophic factor; CNTF; brain-derived neurotrophic factor;
KW BNTF; neurotrophin; NT-3; NT-4; NT-5; neurodegenerative disease;
KW familial dysautonomia; infantile muscular dystrophy;
KW Parkinson's disease; Alzheimer's disease.
OS Homo sapiens.
PN WO9905272-A1.
PD 04-FEB-1999.
PF 27-JUL-1998; G02228.
PR 10-DEC-1997; US-988476.
PR 25-JUL-1997; GB-015823.
PA (UNLO) UNIV COLLEGE LONDON.
PI Latchman DS, Smith MD;
DR WPI; 99-142928/12.
DR N-PSDB; X09010.
PT New polypeptide comprising transcription factor Brn-3a, or its
PT derivative - useful for treating nervous system diseases, preventing
PT cellular apoptosis and increasing nerve regeneration following
PT neuronal damage
PS Disclosure; Page 61-62; 68pp; English.
CC Over expression of transcription factor Brn-3a can protect neuronal
CC cells from apoptosis. Brn-3a also specifically activates expression
CC of the BCl-2 gene in neuronal cells and this activation is mediated
CC via a Brn-3a response element in the 5' regulatory region of the
CC Bcl-2 gene. Both the anti-apoptopic effect of Brn-3a and its ability
CC to activate expression of Bcl-2 are mediated by the N-terminal
CC domain of Brn-3a. Members of the Bcl-2 family perform critical roles
CC in the regulation of selective apoptosis during development of the
CC nervous system. The stimulation of Bcl-2 expression by Brn-3a in a
CC neuron specific manner and consequent protection of neuronal cells
CC from apoptosis suggests that Brn-3a may co-ordinate some aspects of
CC neuronal reorganisation during development or following injury. The
CC elevation of Brn-3a expression by either pharmacological means
CC (compositions comprising one other therapeutic polypeptide e.g.
CC neurotrophic factors, nerve growth factor (NGF), ciliary neurotrophic
CC factor (CNTF), brain-derived neurotrophic factor (BNTF), and
CC neurotrophins NT-3 and NT-4/5) or gene therapy may represent a
CC method for treating human diseases associated with excessive
CC neuronal cell death and/or lack of nerve regeneration, especially
CC neurodegenerative diseases such as familial dysautonomia and
CC infantile muscular dystrophy, and Parkinson's and Alzheimer's
CC disease.
SQ Sequence 423 AA;

W96262 Length: 423 February 11, 2000 15:49 Type: P Check: 3324 ..

1 MMSMNSKQPH FAMHPTLPEH KYPSLHSSSE AIRRACLPTP PLQSNLFASL

51 DETLLARAEA LAAVDIAVSQ GKSHPFKPDA TYHTMNSVPC TSTSTVPLRH

101 HHHHHHHHQA LEPGDLLDHI SSPSLALMAG AGGAGGAGAA AGGGGAHDGP

151 GGGGGPGGGG GPGGGGPGGG GGGGPGGGGG GPGGGLLGGS AHPHPHMHSL

201 GHLSHPAAAA AMNMPSGLPH PGLVAAAAHH GAAAAAAAAA AGQVAAASAA

251 AAVVGAAGLA SICDSDTDPR ELEAFAERFK QRRIKLGVTQ ADVGSALANL

301 KIPGVGSLSQ STICRFESLT LSHNNMIALK PILQAWLEEA EGAQREKMNK

351 PELFNGGEKK RKRTSIAAPE KRSLEAYFAV QPRPSSEKIA AIAEKLDLKK

401 NVVRVWFCNQ RQKQKRMKFS ATY

!!AA_SEQUENCE 1.0
ID Y05204 standard; peptide; 24 AA.
AC Y05204;
DT 17-JUN-1999 (first entry)
DE Human CLIP immunomodulatory peptide.
KW CLIP; immunomodulatory peptide; immune disorder; TH1 cell reduction;
KW MHC class II-associated invariant chain protein; transplant rejection;
KW TH2 cell; immune response; T cell response; T cell activation; arthritis;
KW autoimmune disorder; autoimmune diabetes; multiple sclerosis;
KW systemic lupus erythematosus; cytokine production; human.
OS Homo sapiens.
PN CA2205680-A.
PD 16-NOV-1998.
PF 16-MAY-1997; 205680.
PR 16-MAY-1997; CA-205680.
PA (UYWO-) UNIV WESTERN ONTARIO.
PI Singh B;
DR WPI; 99-244657/21.
PT Use of an MHC class II-associated invariant chain protein (CLIP) for
PT treating or preventing an immune disorder
PS Claim 3; Fig 1; 62pp; English.
CC This sequence represents the human CLIP immunomodulatory peptide.
CC The invention relates to a method of treating or preventing an immune
CC disorder comprises administering an MHC class II-associated invariant
CC chain protein (CLIP) to an affected/susceptible mammal. The CLIP peptides
CC are useful as immunogenic compositions for: (1) increasing the level of
CC TH2 cells; (2) decreasing the level of TH1 cells; (3) preventing or
CC reducing an immune response; (4) reducing or preventing a T cell response
CC to an antigen; (5) reducing or preventing antigen presentation by APCs;
CC and (6) treating or preventing a disorder associated with a T cell
CC response; in mammals, preferably humans. The immune disorders involve
CC increased T cell activation and include transplant rejection and
CC autoimmune disorders, including autoimmune diabetes, arthritis, multiple
CC sclerosis and systemic lupus erythematosus. The CLIP peptides are useful
CC for screening for activity of CLIP fragments/analogues, and for assaying
CC for ability to affect cytokine production such as interleukin-2 or
CC interferon-gamma.
SQ Sequence 24 AA;

Y05204 Length: 24 February 11, 2000 15:49 Type: P Check: 3125 ..

1 KPKPPKPVSK MRMATPLLMQ ALPM

```
!!AA_SEQUENCE 1.0
ID   Y11993 standard; Protein; 102 AA.
AC   Y11993;
DT   18-JUN-1999  (first entry)
DE   Human 5' EST secreted protein SEQ ID No: 593.
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition.
OS   Homo sapiens.
PN   WO9906550-A2.
PD   11-FEB-1999.
PF   31-JUL-1998; IB1232.
PR   01-AUG-1997; US-905144.
PA   (GEST ) GENSET.
PI   Duclert A, Dumas Milne Edwards J, Lacroix B;
DR   WPI; 99-153780/13.
DR   N-PSDB; X40715.
PT   New isolated prostate-derived nucleic acids - used to develop
PT   products which may have cytokine, immune regulatory, haematopoiesis
PT   regulating, anti-inflammatory or tumour inhibition activity
PS   Claim 34; Page 675; 675pp; English.
CC   X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins expressed in prostate, and encode the proteins given in
CC   Y11716 to Y11993 respectively. The proteins given represent the signal
CC   peptide and an N-terminal fragment of a secreted protein. The nucleic
CC   acid sequences can be used for producing secreted human gene products.
CC   They can also be used to develop products for diagnosis and therapy. The
CC   proteins obtained may have cytokine activity, cell proliferation and
CC   differentiation activity, haematopoiesis regulating activity, tissue
CC   growth regulating activity, reproductive hormone regulating activity,
CC   chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC   receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC   activity or other activities. The products can be used in forensic, gene
CC   therapy and chromosome mapping procedures. The sequences can also be used
CC   for obtaining corresponding promoter sequences. The nucleic acids
CC   encoding the signal peptides can be used for directing extracellular
CC   secretion of a polypeptide or the insertion of a polypeptide into a
CC   membrane, or importing a polypeptide into a cell.
SQ   Sequence   102 AA;

 Y11993  Length: 102  February 11, 2000 15:49  Type: P  Check: 4119  ..

       1  MDQFPESVTE NFEYDDLAEA CYIGDIVVFG TVFLSIFYSV IFAIGLVGNL

      51  LVVFALTNSK KPKSVTDIYL LNLALSDLLF VATLPFWTHY LINEKGLHNA

     101  MC

!!AA_SEQUENCE 1.0
ID   Y05545 standard; Protein; 243 AA.
AC   Y05545;
DT   05-JUL-1999  (first entry)
DE   Wheat Type III glutathione transferase subunit ICP/ICV.
KW   Glutathione transferase; GST; glutathione peroxidase; wheat; ICP;
KW   ICV; herbicide resistance; transgenic plant.
OS   Triticum aestivum.
PN   WO9914337-A2.
PD   25-MAR-1999.
PF   16-SEP-1998; G02802.
PR   16-SEP-1997; GB-019727.
PA   (RHON ) RHONE-POULENC AGRIC LTD.
PI   Cole DJ, Cummins I, Edwards R;
DR   WPI; 99-244035/20.
DR   N-PSDB; X25152.
PT   New isolated glutathione transferase subunit polynucleotides
PS   Claim 3; Page 100; 101pp; English.
CC   The present sequence represents ICP and ICV, wheat glutathione
CC   transferase (GST) subunits that resemble the type III GSTs of maize.
CC   ICP/ICV cDNA (see X25152) was isolated from a cDNA library prepared
CC   from fenchlorazole-ethyl (herbicide safener) treated wheat shoots.
CC   The invention provides wheat GST subunits (see Y05537-45) active in
CC   herbicide metabolism.  This is fundamental to understanding GST
CC   detoxification in wheat and in the development of transgenic
CC   herbicide resistant plants expressing wheat GSTs.  The invention
CC   provides methods of identifying compounds capable of metabolism by
CC   GST, or compounds that induce GST expression in graminaceous plants,
CC   and for determining the GST level in a sample of seed or flour.
CC   Transgenic plants, host cells used for production of GST subunits
CC   and GST dimeric proteins, and vectors are also provided.
SQ   Sequence   243 AA;

 Y05545  Length: 243  February 11, 2000 15:49  Type: P  Check: 5919  ..

       1  MAGGEELKLL GWWAPGVSPY VLRAQMALAV KGLSYDYLPE DRWSTSDLLI

      51  ASNPVYKKVP VLIHNGRPVC ESLLILEYLD DAVGLAGNGK PILPADPYSR

     101  AVARFWAAYV NDKLFPSCTG ILKTTKQEER AGKMEETLSG LRHLEAVMAE

     151  CSEGEAEAPF FGGDAIGFLD IALGCYLPWF EAAGRLAGLG PIIDPARTPK

     201  LAAWAERFSV AEPIKALLPG VDKLEEYITT ALYPKWNIAV TGN

!!AA_SEQUENCE 1.0
ID   Y07041 standard; Protein; 207 AA.
AC   Y07041;
DT   02-JUL-1999  (first entry)
DE   Breast cancer associated antigen precursor sequence.
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer.
OS   Homo sapiens.
PN   WO9904265-A2.
PD   28-JAN-1999.
PF   15-JUL-1998; U14679.
PR   22-JUN-1998; US-102322.
PR   17-JUL-1997; US-896164.
PR   10-OCT-1997; US-061599.
PR   10-OCT-1997; US-061765.
PR   10-OCT-1997; US-948705.
PR   11-OCT-1997; GB-021697.
PA   (LUDW-) LUDWIG INST CANCER RES.
PI   Chen Y, Gout I, Gure A, OHare M, Obata Y, Old LJ,
PI   Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E,
PI   Tureci O;
DR   WPI; 99-132448/11.
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
PS   Disclosure; Page 426-427; 787pp; English.
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
SQ   Sequence   207 AA;

 Y07041  Length: 207  February 11, 2000 15:49  Type: P  Check: 1679  ..

       1  MSPLLRRLLL AALLQLAPAQ APVSQPDAPG HQRKVVSWID VYTRATCQPR

      51  EVVVPLTVEL MGTVAKQLVP SCVTVQRCGG CCPDDGLECV PTGQHQVRMQ

     101  ILMIRYPSSQ LGEMSLEEHS QCECRPKKKD SAVKPDRAAT PHHRPQPRSV
```

151 PGWDSAPGAP SPADITHPTP APGPSAHAAP STTSALTPGP AAAAADAAAS

201 SVAKGGA

!!AA_SEQUENCE 1.0
ID Y07007 standard; Protein; 177 AA.
AC Y07007;
DT 02-JUL-1999 (first entry)
DE Breast cancer associated antigen precursor sequence.
KW Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer.
OS Homo sapiens.
PN WO9904265-A2.
PD 28-JAN-1999.
PF 15-JUL-1998; U14679.
PR 22-JUN-1998; US-102322.
PR 17-JUL-1997; US-896164.
PR 10-OCT-1997; US-061599.
PR 10-OCT-1997; US-061765.
PR 10-OCT-1997; US-948705.
PR 11-OCT-1997; GB-021697.
PA (LUDW-) LUDWIG INST CANCER RES.
PI Chen Y, Gout I, Gure A, OHare M, Obata Y, Old LJ,
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E,
PI Tureci O;
DR WPI; 99-132448/11.
PT New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers
PS Disclosure; Page 391-392; 787pp; English.
CC The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer.
SQ Sequence 177 AA;

Y07007 Length: 177 February 11, 2000 15:49 Type: P Check: 4692 ..

1 LCCMHYCCKS CWNEYLTTRI EQNLVLNCTC PIADCPAQPT GAFIRAIVSS

51 PEVISKYKAL LRGYVESCSN LTWCTNPQGC DRILCRQGLG CGTTCSKCGW

101 ASCFNCSFPE AHYPASCGHM SQWVDDGGYY DGMSVEAKHL AKLISKRCPS

151 CQAPIENEGC LHMTCAKCNH GFCWRCL

!!AA_SEQUENCE 1.0
ID W93953 standard; Protein; 352 AA.
AC W93953;
DT 30-JUN-1999 (first entry)
DE Human regulatory molecule HRM-9 protein.
KW Human regulatory molecule; HRM-9; cytostatic activity; immune modulator;
KW transcription factor; enhancer; cell proliferation stimulation; cancer;
KW treatment; microarray; detection; diagnosis; cell proliferation disease;
KW leukemia; lymphoma; myeloma; adenocarcinoma; sarcoma; bladder; bone;
KW brain; lung; liver; ovary; skin; teratocarcinoma; immune response;
KW allergy; asthma; diabetes; multiple sclerosis; Grave's disease;
KW myasthenia gravis.
OS Homo sapiens.
PN WO9915658-A2.
PD 01-APR-1999.
PF 22-SEP-1998; U19839.
PR 23-SEP-1997; US-933750.
PA (INCY-) INCYTE PHARM INC.
PI Au-Young J, Bandman O, Guegler KJ, Hillman JL, Lal P,
PI Shah P, Yue H;
DR WPI; 99-254710/21.
DR N-PSDB; X24067.
PT New human regulatory molecules
PS Claim 1; Page 75; 76pp; English.
CC This invention describes novel human regulatory molecules (HRM) which
CC have cytostatic activity and act as immune modulators, transcription
CC factors or enhancers. The HRMs can be used to stimulate cell
CC proliferation. Antagonists and agonists of the proteins of the invention
CC can be used to treat cancer. The encoding nucleic acids can be used in
CC microarrays to detect polynucleotides (and their expression levels) that
CC encode HRMs in a biological sample. The HRMs and microarrays can be used
CC diagnose, treat or prevent cell proliferation diseases especially cancer,
CC e.g. leukemia, lymphoma, myeloma, adenocarcinoma, sarcoma, cancer of e.g.
CC bladder, bone, brain, lung, liver, ovary, skin, etc, teratocarcinoma, or
CC to treat or prevent immune responses e.g. allergies, asthma, diabetes,
CC multiple sclerosis, Grave's disease or myasthenia gravis.
SQ Sequence 352 AA;

W93953 Length: 352 February 11, 2000 15:49 Type: P Check: 7759 ..

1 MHVVAPASLR LGTGTNLPPS PTCLTKLALP PAAEPSLLAM SQSRHRAEAP

51 PLEREDSGTF SLGKMITAKP GKTPIQVLHE YGMKTKNIPV YECERSDVQI

101 HVPTFTFRVT VGDITCTGEG TSKKLAKHRA AEAAINILKA NASICFAVPD

151 PLMPDPSKQP KNQLNPIGSL QELAIHHGWR LPEYTLSQEG GPAHKREYTT

201 ICRLESFMET GKGASKKQAK RNAAEKFLAK FSNISPENHI SLTNVVGHSL

251 GCTWHSLRNS PGEKINLLKR SLLSIPNTDY IQLLSEIAKE QGFNITYLDI

301 DELSANGQYQ CLAELSTSPI TVCHGSGISC GNAQSDAAHN ALQYLKIIAE

351 RK

!!AA_SEQUENCE 1.0
ID Y04959 standard; Protein; 273 AA.
AC Y04959;
DT 06-JUL-1999 (first entry)
DE Mycobacterium species protein sequence 42D.
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection.
OS Mycobacterium sp.
PN WO9909186-A2.
PD 25-FEB-1999.
PF 14-AUG-1998; F01813.
PR 11-SEP-1997; FR-011325.
PR 14-AUG-1997; FR-010404.
PA (INSP) INST PASTEUR.
PI Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI Guigueno A;
DR WPI; 99-181045/15.
DR N-PSDB; X34210.
PT Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PS Claim 32; Fig 42D; 309pp; French.
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
SQ Sequence 273 AA;

Y04959 Length: 273 February 11, 2000 15:49 Type: P Check: 8454 ..

```
       1  MANKRGNAGQ PLPLSDRDDD HMQGHWLLAR LGKRVLRPGG VELTRTLLAR

      51  AEVTDADVLE LAPGLGRTAA EILARNPRSY VGAESDPNAA NLVRHVLAGR

     101  GDVRVTDAAD TGLSDASADV VIGEAMLTMQ GNAAKHTIVA EAARVLRPGG

     151  RYAIHELALV PDDVAEQVRT DLRQSLARAL KVNARPLTVA EWSHLLAGHG

     201  LVVEHVVTAS MALLQPRRVI ADEGLLGALR FAGNLLIHRA ARRRVLLMRH

     251  TFRRHRERLT AVAIVAHKPH VDS
```

```
!!AA_SEQUENCE 1.0
ID   Y04960 standard; Protein; 280 AA.
AC   Y04960;
DT   06-JUL-1999  (first entry)
DE   Mycobacterium species protein sequence 42F.
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection.
OS   Mycobacterium sp.
PN   WO9909186-A2.
PD   25-FEB-1999.
PF   14-AUG-1998; F01813.
PR   11-SEP-1997; FR-011325.
PR   14-AUG-1997; FR-010404.
PA   (INSP ) INST PASTEUR.
PI   Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI   Guigueno A;
DR   WPI; 99-181045/15.
DR   N-PSDB; X34211.
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in
PT   infection-associated protein expression
PS   Claim 32; Fig 42F; 309pp; French.
CC   Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC   from various Mycobacterium species microorganisms.  The encoding
CC   nucleotide sequences can be used as primers and probes for methods
CC   for detecting and identifying mycobacteria, especially belonging to
CC   the M. tuberculosis complex.  The encoded proteins can be used in
CC   vaccines for immunisation against a bacterial or viral infection.
SQ   Sequence   280 AA;
```

```
 Y04960  Length: 280  February 11, 2000 15:49  Type: P  Check: 8320  ..

       1  KPAEVGSMAN KRGNAGQPLP LSDRDDDHMQ GHWLLARLGK RVLRPGGVEL

      51  TRTLLARAEV TDADVLELAP GLGRTAAEIL ARNPRSYVGA ESDPNAANLV

     101  RHVLAGRGDV RVTDAADTGL SDASADVVIG EAMLTMQGNA AKHTIVAEAA

     151  RVLRPGGRYA IHELALVPDD VAEQVRTDLR QSLARALKVN ARPLTVAEWS

     201  HLLAGHGLVV EHVVTASMAL LQPRRVIADE GLLGALRFAG NLLIHRAARR

     251  RVLLMRHTFR RHRERLTAVA IVAHKPHVDS
```

```
!!AA_SEQUENCE 1.0
ID   Y04901 standard; Protein; 72 AA.
AC   Y04901;
DT   06-JUL-1999  (first entry)
DE   Mycobacterium species protein sequence 31C.
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection.
OS   Mycobacterium sp.
PN   WO9909186-A2.
PD   25-FEB-1999.
PF   14-AUG-1998; F01813.
PR   11-SEP-1997; FR-011325.
PR   14-AUG-1997; FR-010404.
PA   (INSP ) INST PASTEUR.
PI   Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI   Guigueno A;
DR   WPI; 99-181045/15.
DR   N-PSDB; X34154.
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in
PT   infection-associated protein expression
PS   Claim 32; Fig 31C; 309pp; French.
CC   Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC   from various Mycobacterium species microorganisms.  The encoding
CC   nucleotide sequences can be used as primers and probes for methods
CC   for detecting and identifying mycobacteria, especially belonging to
CC   the M. tuberculosis complex.  The encoded proteins can be used in
CC   vaccines for immunisation against a bacterial or viral infection.
SQ   Sequence   72 AA;
```

```
 Y04901  Length: 72  February 11, 2000 15:49  Type: P  Check: 6286  ..

       1  ANTREGSATA ALPAPWLLNN TSQPRTLPVC GRINDPNSTN TRIATTKALA

      51  PGSISRATRR SAASISARRA SR
```

```
!!AA_SEQUENCE 1.0
ID   Y04743 standard; Protein; 110 AA.
AC   Y04743;
DT   06-JUL-1999  (first entry)
DE   Mycobacterium species protein sequence 1#2.
KW   Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW   hybridisation; detection; vaccine; immunisation; infection.
OS   Mycobacterium sp.
PN   WO9909186-A2.
PD   25-FEB-1999.
PF   14-AUG-1998; F01813.
PR   11-SEP-1997; FR-011325.
PR   14-AUG-1997; FR-010404.
PA   (INSP ) INST PASTEUR.
PI   Gicquel B, Lim EM, Pelicic V, Portnoi D, Goguet de la Salmoniere Y,
PI   Guigueno A;
DR   WPI; 99-181045/15.
DR   N-PSDB; X34001.
PT   Mycobacterial DNA vectors containing reporter constructs - for
PT   identifying coding or promoter sequences involved in
PT   infection-associated protein expression
PS   Claim 32; Fig 1; 309pp; French.
CC   Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC   from various Mycobacterium species microorganisms.  The encoding
CC   nucleotide sequences can be used as primers and probes for methods
CC   for detecting and identifying mycobacteria, especially belonging to
CC   the M. tuberculosis complex.  The encoded proteins can be used in
CC   vaccines for immunisation against a bacterial or viral infection.
SQ   Sequence   110 AA;
```

```
 Y04743  Length: 110  February 11, 2000 15:49  Type: P  Check: 5429  ..

       1  MAGDQELELR FDVPLYTLAE ASRYLVVPRA TLATWADGYE RRPANAPAVQ

      51  GQPIAFDAYS VAQLFGDVTG ARVAGVQPQR HHIRPVRLRG PLGGVGCLRH

     101  PRQFAGYLSQ
```

```
!!AA_SEQUENCE 1.0
ID   Y07239 standard; Protein; 141 AA.
AC   Y07239;
DT   06-JUL-1999  (first entry)
DE   Fragment of human placental STAT3 protein.
KW   Placenta; isoform; human; STAT3; intracellular; transcription factor;
KW   Signal Transducer and Activator of Transcription; allele; growth arrest;
KW   hepatic acute-phase protein; monocytic cell; myeloma; autoimmune disease;
KW   inflammation.
OS   Homo sapiens.
PN   EP-906953-A1.
PD   07-APR-1999.
PF   16-SEP-1997; 116061.
PR   16-SEP-1997; EP-116061.
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PI   Della Pietra L, Serlupi-Crescenzi O;
```

DR WPI; 99-207107/18.
DR N-PSDB; X29975.
PT New allelic variant of human STAT3 useful in treating autoimmune or
PT inflammatory diseases
PS Disclosure; Page 8-9; 22pp; English.
CC This sequence represents a fragment of the placental isoform of human
CC Signal Transducer and Activator of Transcription (STAT3) intracellular
CC transcription factor (Akira et al., Cell 77, 63-71 (1994)). The
CC invention relates to isolation of allelic variants of the placental
CC hSTAT3 sequence. hSTAT3 plays a role in the upregulation of hepatic
CC acute-phase proteins, growth arrest of monocytic cells and in the
CC survival of myeloma cells and so may be used to treat or diagnose
CC autoimmune or inflammatory diseases.
SQ Sequence 141 AA;

Y07239 Length: 141 February 11, 2000 15:49 Type: P Check: 2144 ..

1 WLDNIIDLVK KYILALWNEG YIMGFISKER ERAILSTKPP GTFLLRFSES

51 SKEGGVTFTW VEKDISGKTQ IQSVEPYTKQ QLNNMSFAEI IMGYKIMDAT

101 NILLSPLVYL YPDIPKEEAF GKYCRPESQE HPEADPGSAA P

!!AA_SEQUENCE 1.0
ID Y07240 standard; Protein; 141 AA.
AC Y07240;
DT 06-JUL-1999 (first entry)
DE Fragment of human hepatic STAT3 protein.
KW Placenta; isoform; human; STAT3; intracellular; transcription factor;
KW Signal Transducer and Activator of Transcription; allele; growth arrest;
KW hepatic acute-phase protein; monocytic cell; myeloma; autoimmune disease;
KW inflammation.
OS Homo sapiens.
PN EP-906953-A1.
PD 07-APR-1999.
PF 16-SEP-1997; 116061.
PR 16-SEP-1997; EP-116061.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI Della Pietra L, Serlupi-Crescenzi O;
DR WPI; 99-207107/18.
DR N-PSDB; X29976.
PT New allelic variant of human STAT3 useful in treating autoimmune or
PT inflammatory diseases
PS Claim 2; Page 10-11; 22pp; English.
CC This sequence represents a fragment of the hepatic allelic isoform
CC of human Signal Transducer and Activator of Transcription (STAT3)
CC intracellular transcription factor (Akira et al., Cell 77, 63-71
CC (1994)). The invention relates to isolation of allelic variants of
CC the placental hSTAT3 sequence. hSTAT3 plays a role in the upregulation
CC of hepatic acute-phase proteins, growth arrest of monocytic cells and
CC in the survival of myeloma cells and so may be used to treat or diagnose
CC autoimmune or inflammatory diseases.
SQ Sequence 141 AA;

Y07240 Length: 141 February 11, 2000 15:49 Type: P Check: 2614 ..

1 WLDNIIDLVK KYILALWNEG YIMGFISKER ERAILSTKPP GTFLLRFSES

51 SKEGGVTFTW VEKDISGKTQ IQSVEPYTKQ QLNNMSFAEI IMGYKIMDAT

101 NILVSPLVYL YPDIPKEEAF GKYCRPESQE HPEADPGSAA P

!!AA_SEQUENCE 1.0
ID Y07241 standard; Protein; 141 AA.
AC Y07241;
DT 06-JUL-1999 (first entry)
DE Fragment of mouse hepatic STAT3 protein.
KW Placenta; isoform; human; STAT3; intracellular; transcription factor;
KW Signal Transducer and Activator of Transcription; allele; growth arrest;
KW hepatic acute-phase protein; monocytic cell; myeloma; autoimmune disease;
KW inflammation.
OS Mus sp.
PN EP-906953-A1.
PD 07-APR-1999.
PF 16-SEP-1997; 116061.
PR 16-SEP-1997; EP-116061.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI Della Pietra L, Serlupi-Crescenzi O;
DR WPI; 99-207107/18.
DR N-PSDB; X29977.
PT New allelic variant of human STAT3 useful in treating autoimmune or
PT inflammatory diseases
PS Disclosure; Page 12-13; 22pp; English.
CC This sequence represents a fragment of the hepatic allelic isoform
CC of mouse Signal Transducer and Activator of Transcription (STAT3)
CC intracellular transcription factor (Akira et al., Cell 77, 63-71
CC (1994)). The invention relates to isolation of allelic variants of
CC the placental human STAT3 sequence. hSTAT3 plays a role in the
CC upregulation of hepatic acute-phase proteins, growth arrest of monocytic
CC cells and in the survival of myeloma cells and so may be used to treat
CC or diagnose autoimmune or inflammatory diseases.
SQ Sequence 141 AA;

Y07241 Length: 141 February 11, 2000 15:49 Type: P Check: 3064 ..

1 WLDNIIDLVK KYILALWNEG YIMGFISKER ERAILSTKPP GTFLLRFSES

51 SKEGGVTFTW VEKDISGKTQ IQSVEPYTKQ QLNNMSFAEI IMGYKIMDAT

101 NILVSPLVYL YPDIPKEEAF GKYCRPESQE HPEADPGSSA P